

LinearRegressionGenderScopus2minimalRobustLoop.R

MikeT

Wed Jun 12 11:48:25 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

##
## 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 #Allow the extra time for all fields,
just in case
#FieldNeedsExtraTime <- c(1311, 3100, 2700, 2709, 2714, 2719,2722, 2734,
2744, 2747, 3104, 3107, 1701, 1706, 3302,3303,3309, 3320, 3322) #Fields known
to need extra time
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$LastAuthorFemale > -1,] #Female 1, Male 0, Unknown -1
AllScopusDataOlderFirstLastGendered$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$YMresiduals <- resid(YearLm)
    YearGenderLm <- lm(YMresiduals~FirstAuthorFemale,
data=AllScopusDataOlderFirstGendered)
    print(bartlett.test(YMresiduals~FirstAuthorFemale,
data=AllScopusDataOlderFirstGendered)) #Homogeneity of Variances test - this
is the key test - should not fail by much, but might becuae older years have
fewer females and higher variance
    plot(YearGenderLm, which = 1) #Residuals vs. fitted. Check that width
of residual bar does not vary too much for homoscedasticity; look out for
cone shape
    #If problems here, check the outliers in regression 1
  }, error = function(e) return("failed Heteroscedasticity checks"))
  tryCatch({
    #####
    # Test for different team sizes for male and female authors - first and
last gendered
    #####
    #First author gender 2018
    GenderAuthors2018 <-
split(AllYearsBothGendered2018$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$UniqueAuthors,AllYearsBothGendered2018$LastAuthorFemale)
  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$UniqueAuthors[i])
  }

AllScopusDataOlderFirstLastGendered$UniqueAuthors=factor(AllScopusDataOlderFirstLastGendered$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,

```

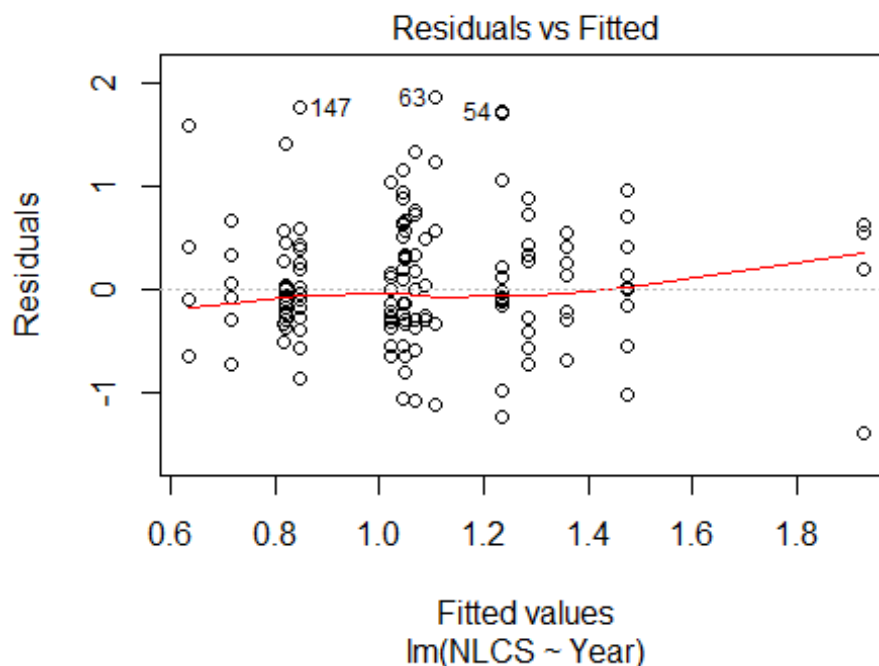
```

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
## 19 25 19 25 26 38 31 28 17 9 19 13 12 17 8
## 2011 2012
## 8 16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 14 7 11 10 18 17 13 10 4 10 8 6 10 4
## 2011 2012
## 4 5
##

```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6   14    7   11    9   17   15   13    9    3    9    6    6   10    4
## 2011 2012
##      4    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 27, df = 16, p-value = 0.04
```

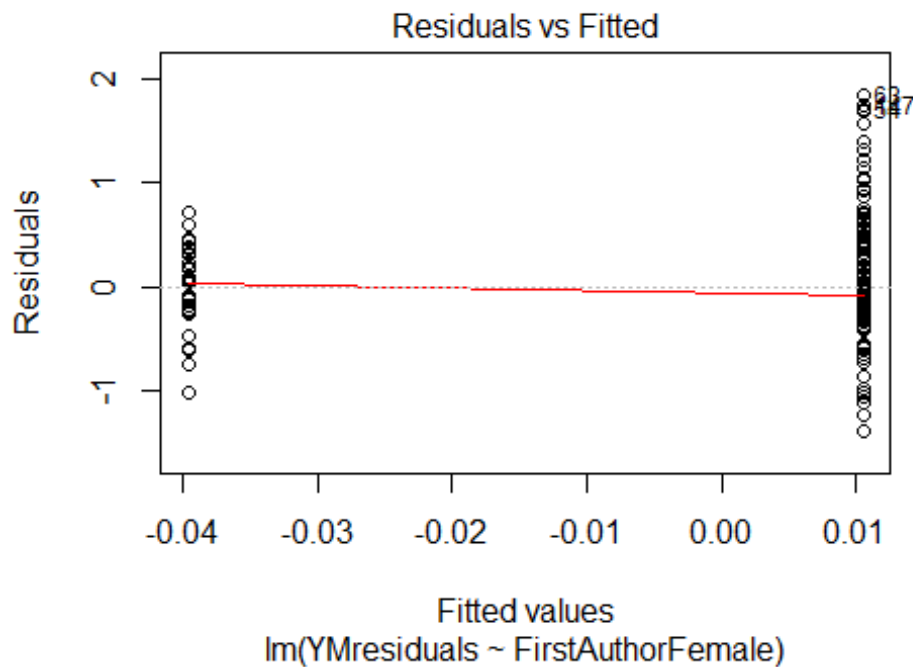


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 7.9, df = 1, p-value = 0.005
## [1] "Female first author team size 2018 geometric mean: 4.99060584155396"
## [1] "Male first author team size 2018 geometric mean: 3.22230618444051"
## 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 = 270, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.91486764116886"
## [1] "Male last author team size 2018 geometric mean: 3.69289194076225"

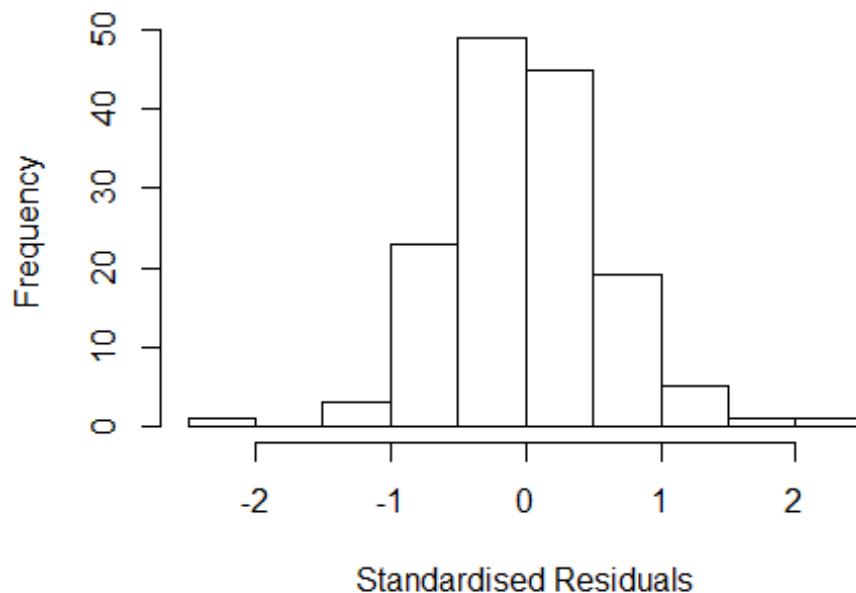
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190, 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	4.303	1	2.074
LastAuthorFemale	2.369	1	1.539
UniqueAuthors	17.903	4	1.434
Year	47.177	16	1.128

Residuals from first and last author and team size



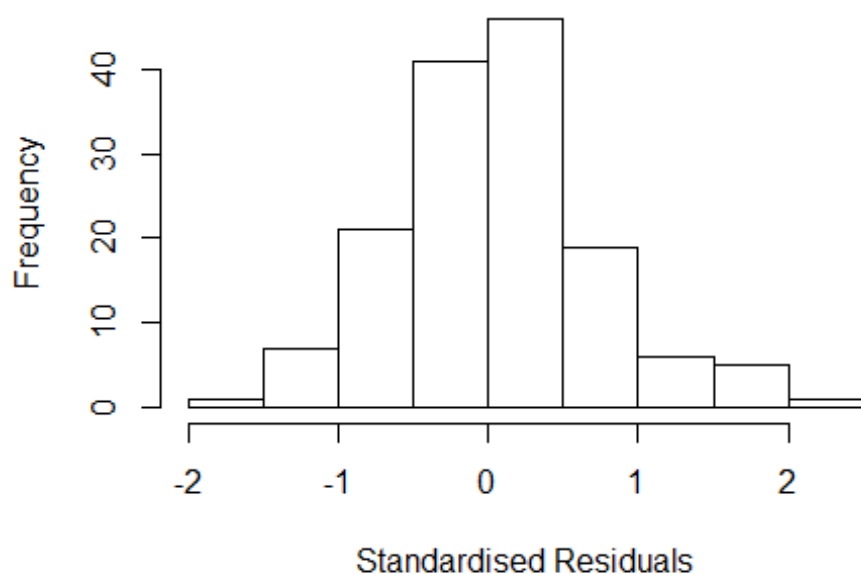
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -2.1317 -0.3645 -0.0481 0.3785 2.4032
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.32068 0.30409 1.05 0.29368
## FirstAuthorFemale1 -0.25158 0.15723 -1.60 0.11213
## LastAuthorFemale1 -0.02102 0.13084 -0.16 0.87261
## UniqueAuthors2 0.59107 0.17234 3.43 0.00082 ***
## UniqueAuthors3 0.78798 0.19311 4.08 8.0e-05 ***
## UniqueAuthors4 0.68573 0.21396 3.20 0.00172 **
## UniqueAuthors5 0.87503 0.17251 5.07 1.4e-06 ***
## Year1997 0.53930 0.37681 1.43 0.15488
## Year1998 0.24115 0.48182 0.50 0.61762
## Year1999 0.65616 0.35456 1.85 0.06660 .
```

```

## Year2000      0.34082      0.39299      0.87  0.38749
## Year2001      0.28022      0.33537      0.84  0.40502
## Year2002     -0.15707      0.31796     -0.49  0.62219
## Year2003      0.31031      0.34654      0.90  0.37228
## Year2004      0.57611      0.32927      1.75  0.08265 .
## Year2005      0.41358      0.33310      1.24  0.21673
## Year2006      0.25684      0.33099      0.78  0.43924
## Year2007      0.45970      0.36103      1.27  0.20530
## Year2008     -0.00328      0.35894     -0.01  0.99273
## Year2009      0.70377      0.34478      2.04  0.04335 *
## Year2010      0.25046      0.37068      0.68  0.50050
## Year2011      1.57506      0.65095      2.42  0.01699 *
## Year2012      0.10409      0.38317      0.27  0.78633
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.557
## Multiple R-squared:  0.405, Adjusted R-squared:  0.299
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 140 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0229 0.8960 0.9520 0.9030 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      6.80e-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 2.179 1      1.476
## LastAuthorFemale 1.883 1      1.372
## Year      4.097 16      1.045

```

Residuals from first and last author



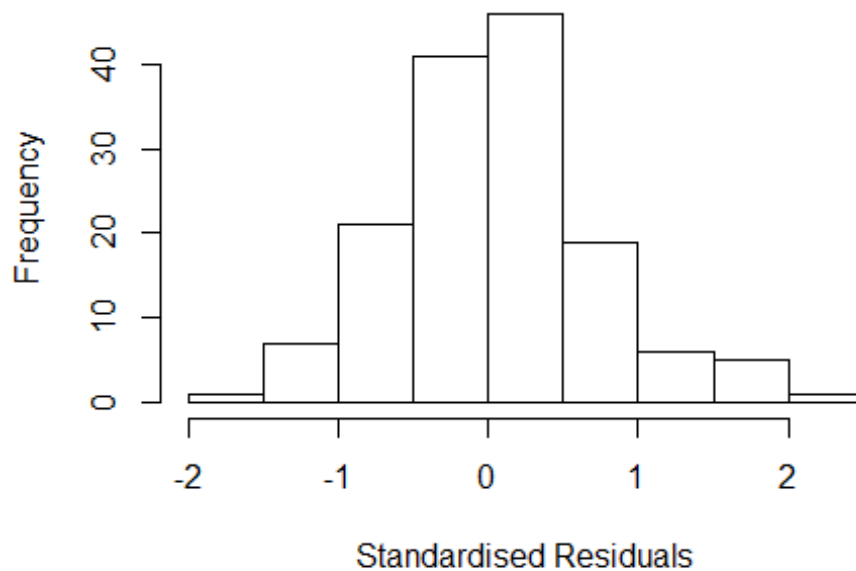
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5989 -0.3494 0.0118 0.3809 2.2251
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4636 0.3243 1.43 0.1553
## FirstAuthorFemale1 -0.0476 0.1327 -0.36 0.7203
## LastAuthorFemale1 0.0230 0.1269 0.18 0.8567
## Year1997 0.6344 0.4110 1.54 0.1252
## Year1998 0.2763 0.8653 0.32 0.7500
## Year1999 0.5899 0.4061 1.45 0.1488
## Year2000 0.6696 0.4276 1.57 0.1198
## Year2001 0.3325 0.3398 0.98 0.3296
## Year2002 0.3227 0.3288 0.98 0.3283
## Year2003 0.5404 0.3442 1.57 0.1188
## Year2004 0.8988 0.3873 2.32 0.0219 *
## Year2005 0.7223 0.3760 1.92 0.0569 .
```

```

## Year2006          0.5584      0.3692      1.51      0.1329
## Year2007          0.9169      0.3764      2.44      0.0162 *
## Year2008          0.2672      0.3729      0.72      0.4749
## Year2009          1.0324      0.3909      2.64      0.0093 **
## Year2010          0.3873      0.4002      0.97      0.3350
## Year2011          1.6873      0.5567      3.03      0.0030 **
## Year2012          0.3389      0.6952      0.49      0.6267
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.617
## Multiple R-squared:  0.194, Adjusted R-squared:  0.081
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 128 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.165  0.860  0.949  0.889  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      6.80e-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 2.143 1      1.464
## Year              2.143 16      1.024

```

Residuals from first author



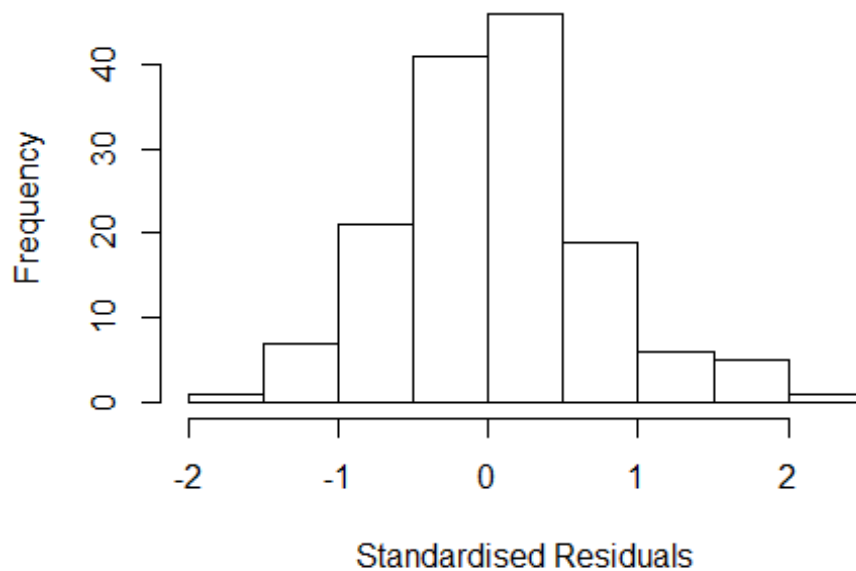
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5901 -0.3529 0.0198 0.3797 2.2124
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4712 0.3208 1.47 0.1444
## FirstAuthorFemale1 -0.0422 0.1347 -0.31 0.7546
## Year1997 0.6323 0.4108 1.54 0.1262
## Year1998 0.2814 0.8543 0.33 0.7424
## Year1999 0.5843 0.4019 1.45 0.1484
## Year2000 0.6629 0.4262 1.56 0.1223
## Year2001 0.3280 0.3381 0.97 0.3338
## Year2002 0.3272 0.3249 1.01 0.3157
## Year2003 0.5332 0.3416 1.56 0.1210
## Year2004 0.8960 0.3747 2.39 0.0182 *
## Year2005 0.7128 0.3721 1.92 0.0576 .
## Year2006 0.5526 0.3647 1.52 0.1322
```

```

## Year2007          0.9156      0.3742      2.45      0.0157 *
## Year2008          0.2610      0.3702      0.70      0.4822
## Year2009          1.0302      0.3879      2.66      0.0089 **
## Year2010          0.3758      0.3968      0.95      0.3453
## Year2011          1.6709      0.5494      3.04      0.0029 **
## Year2012          0.3365      0.6788      0.50      0.6210
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.622
## Multiple R-squared:  0.192, Adjusted R-squared:  0.0852
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.179  0.857   0.947   0.889   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      6.80e-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 2.034 1      1.426
## Year              2.034 16      1.022

```


Residuals from last author



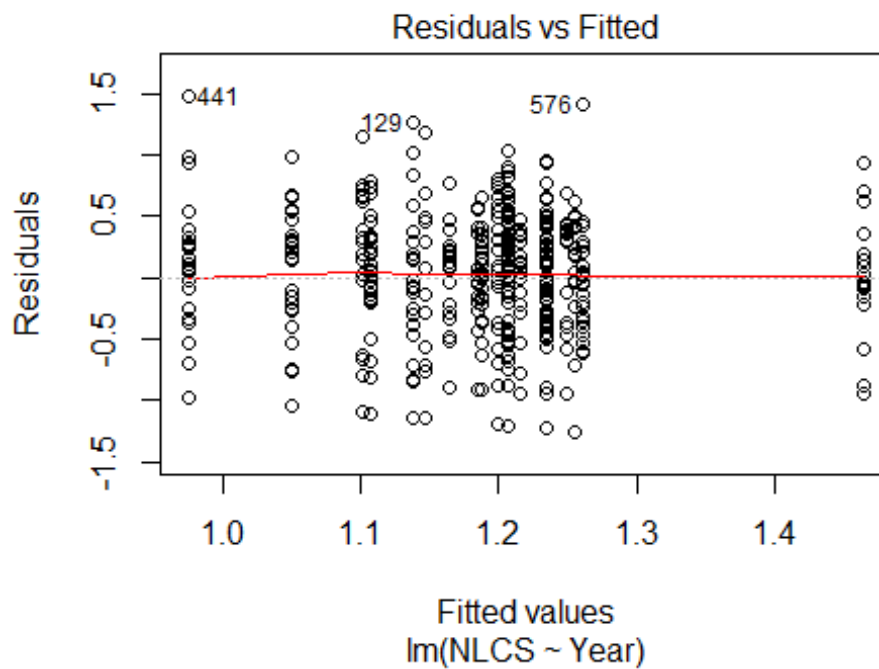
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5577 -0.3398  0.0123  0.3904  2.2217
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.45950    0.32620   1.41   0.1613
## LastAuthorFemale1 0.00811    0.13610   0.06   0.9526
## Year1997        0.63687    0.41342   1.54   0.1259
## Year1998        0.28375    0.82977   0.34   0.7329
## Year1999        0.59155    0.41538   1.42   0.1568
## Year2000        0.66712    0.43235   1.54   0.1253
## Year2001        0.32973    0.34229   0.96   0.3372
## Year2002        0.33152    0.32954   1.01   0.3163
## Year2003        0.52597    0.34587   1.52   0.1308
## Year2004        0.90598    0.37556   2.41   0.0173 *
## Year2005        0.70933    0.37678   1.88   0.0620 .
## Year2006        0.55424    0.37341   1.48   0.1402
```

```

## Year2007      0.90901    0.37774    2.41    0.0175 *
## Year2008      0.25854    0.37491    0.69    0.4917
## Year2009      1.01269    0.37755    2.68    0.0083 **
## Year2010      0.35646    0.39996    0.89    0.3745
## Year2011      1.65015    0.53892    3.06    0.0027 **
## Year2012      0.34634    0.67747    0.51    0.6101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.622
## Multiple R-squared:  0.191, Adjusted R-squared:  0.0842
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 129 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.175  0.860  0.951  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      6.80e-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: 147"
## [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
##   40   50   62   40   37   60   41   41   27   27   42   35   39   47   41
## 2011 2012
##   77   94
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   27   18   23   22   13   15   15   29   17   17   26   22   27   31   26
## 2011 2012

```

```
## 60 73
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 26 17 17 19 13 13 12 26 15 15 20 18 23 26 25
## 2011 2012
## 56 69
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 26, df = 16, p-value = 0.05
```



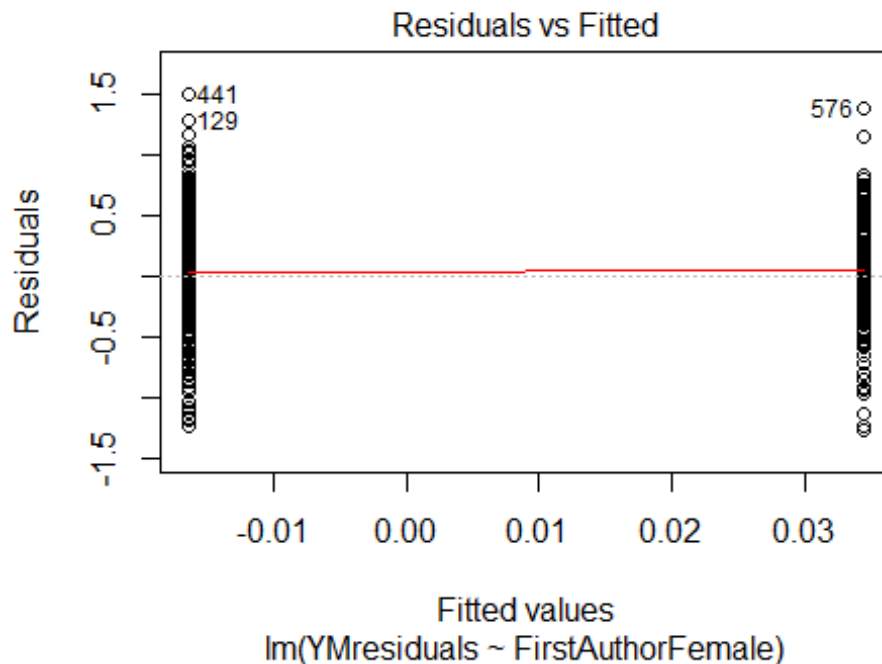
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.86, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 4.32362067735943"
## [1] "Male first author team size 2018 geometric mean: 3.16707631597405"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

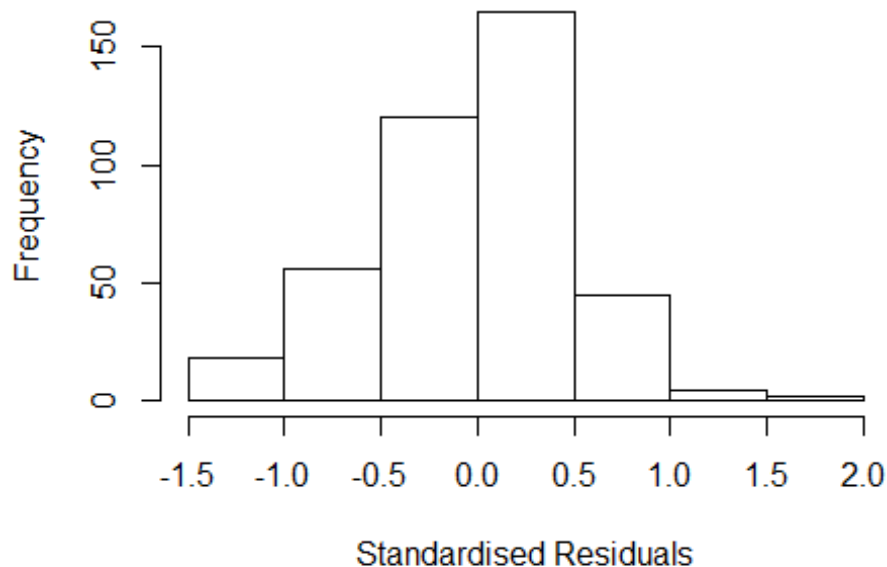
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 590, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.47253802127645"
## [1] "Male last author team size 2018 geometric mean: 3.55043121170024"

## 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 = 480, 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 1.515  1      1.231
## LastAuthorFemale  1.349  1      1.162
## UniqueAuthors    2.518  4      1.122
## Year              3.355 16      1.039
```

Residuals from first and last author and team size



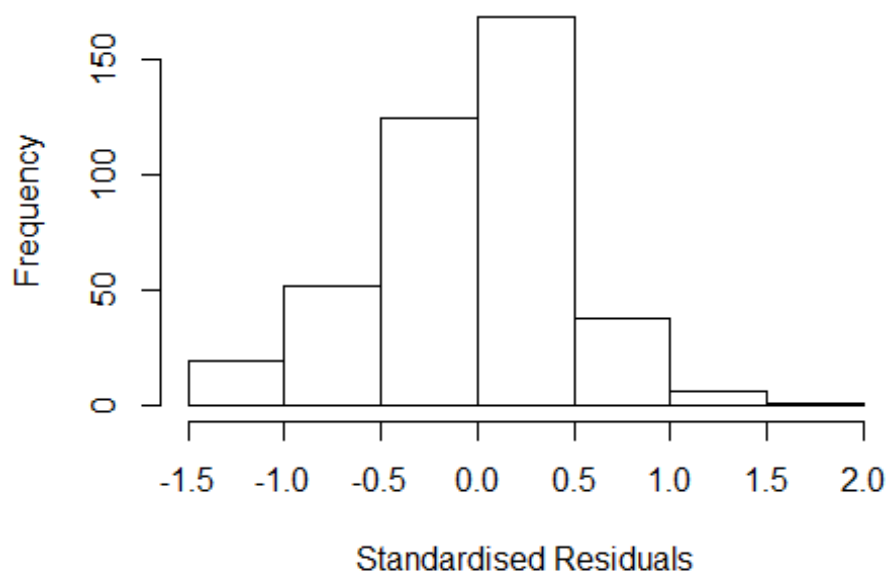
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3916 -0.2824 0.0295 0.2808 1.7398
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1225 0.1397 8.04 1.1e-14 ***
## FirstAuthorFemale1 0.0641 0.0573 1.12 0.26434
## LastAuthorFemale1 -0.0435 0.0712 -0.61 0.54193
## UniqueAuthors2 0.2512 0.0937 2.68 0.00767 **
## UniqueAuthors3 0.1948 0.0977 1.99 0.04699 *
## UniqueAuthors4 0.3042 0.0997 3.05 0.00244 **
## UniqueAuthors5 0.3341 0.0972 3.44 0.00065 ***
## Year1997 0.1951 0.1902 1.03 0.30566
## Year1998 -0.0941 0.2052 -0.46 0.64697
## Year1999 -0.0351 0.1968 -0.18 0.85848
```

```

## Year2000      -0.2122      0.2616      -0.81      0.41773
## Year2001      -0.0615      0.1695      -0.36      0.71681
## Year2002      -0.1013      0.1542      -0.66      0.51162
## Year2003      -0.4143      0.1687      -2.46      0.01451 *
## Year2004       0.0361      0.1809       0.20      0.84200
## Year2005      -0.1605      0.1622      -0.99      0.32313
## Year2006      -0.1226      0.1576      -0.78      0.43703
## Year2007      -0.0759      0.1634      -0.46      0.64270
## Year2008      -0.2077      0.1468      -1.41      0.15802
## Year2009      -0.1641      0.1514      -1.08      0.27915
## Year2010      -0.2778      0.1690      -1.64      0.10101
## Year2011      -0.1200      0.1407      -0.85      0.39428
## Year2012      -0.1496      0.1446      -1.04      0.30126
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0618
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 376 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0837 0.8450 0.9550 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          2.44e-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 1.451 1      1.205
## LastAuthorFemale  1.245 1      1.116
## Year              1.530 16      1.013

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.3669 -0.3071 0.0379 0.3234 1.5779
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2488 0.1466 8.52 3.6e-16 ***
## FirstAuthorFemale1 0.0853 0.0568 1.50 0.134
## LastAuthorFemale1 -0.0315 0.0684 -0.46 0.645
## Year1997 0.2322 0.1898 1.22 0.222
## Year1998 0.0171 0.2078 0.08 0.934
## Year1999 0.0480 0.2069 0.23 0.817
## Year2000 -0.0994 0.2727 -0.36 0.716
## Year2001 -0.0285 0.2048 -0.14 0.889
## Year2002 -0.0760 0.1700 -0.45 0.655
## Year2003 -0.3787 0.1909 -1.98 0.048 *
## Year2004 0.1182 0.1918 0.62 0.538
## Year2005 -0.0910 0.1822 -0.50 0.618
```

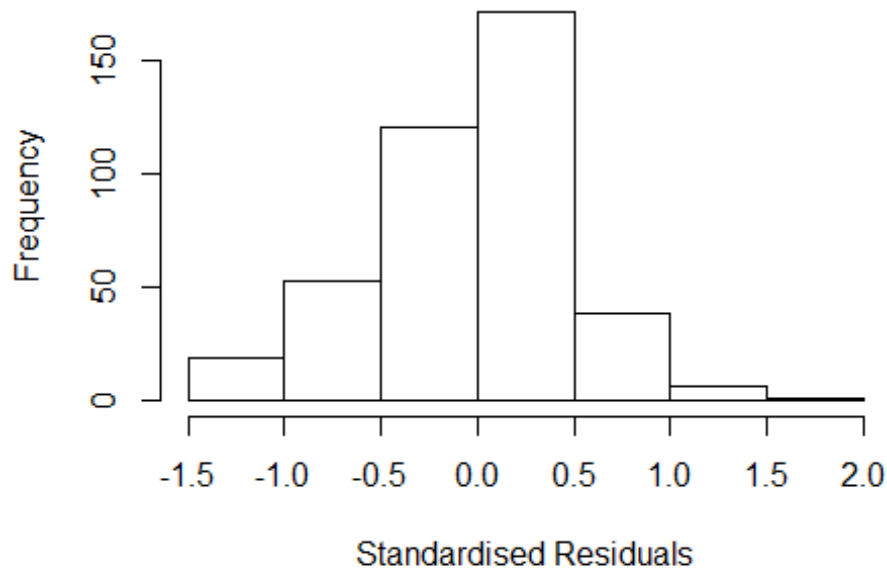


```

## Year2006          -0.0086      0.1701   -0.05    0.960
## Year2007           0.0198      0.1650    0.12    0.905
## Year2008          -0.1190      0.1626   -0.73    0.465
## Year2009          -0.0866      0.1701   -0.51    0.611
## Year2010          -0.1721      0.1864   -0.92    0.356
## Year2011          -0.0188      0.1575   -0.12    0.905
## Year2012          -0.0236      0.1573   -0.15    0.881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.0581, Adjusted R-squared:  0.0148
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 381 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##   0.213  0.866  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      2.44e-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 1.322 1      1.150
## Year              1.322 16      1.009

```

Residuals from first author



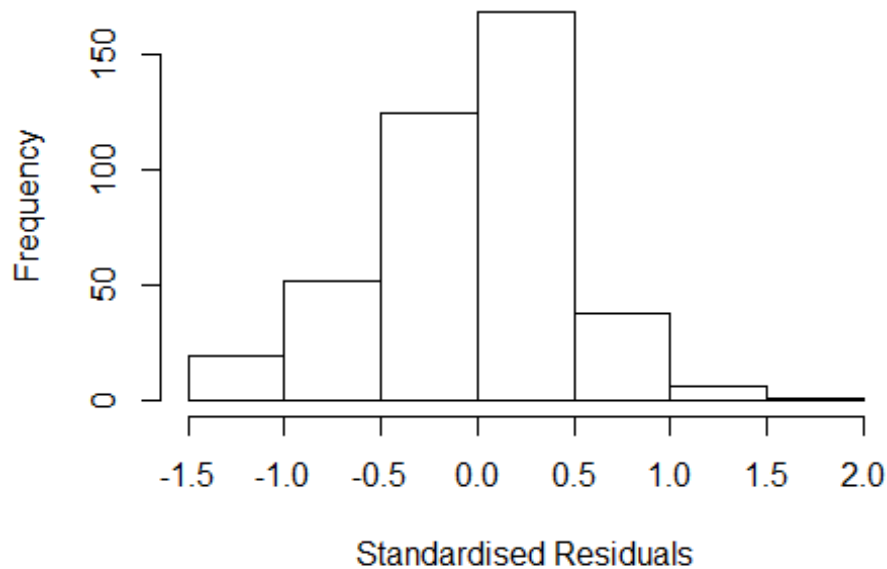
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3628 -0.3047 0.0393 0.3290 1.5765
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.24614 0.14527 8.58 2.3e-16 ***
## FirstAuthorFemale1 0.07684 0.05460 1.41 0.160
## Year1997 0.23116 0.18896 1.22 0.222
## Year1998 0.01747 0.20689 0.08 0.933
## Year1999 0.05071 0.20637 0.25 0.806
## Year2000 -0.09824 0.27150 -0.36 0.718
## Year2001 -0.02986 0.20260 -0.15 0.883
## Year2002 -0.08023 0.16921 -0.47 0.636
## Year2003 -0.37461 0.18926 -1.98 0.048 *
## Year2004 0.11667 0.19018 0.61 0.540
## Year2005 -0.09638 0.18099 -0.53 0.595
## Year2006 -0.00594 0.16892 -0.04 0.972
```

```

## Year2007          0.01671    0.16347    0.10    0.919
## Year2008          -0.11937    0.16169   -0.74    0.461
## Year2009          -0.09079    0.16814   -0.54    0.590
## Year2010          -0.17218    0.18565   -0.93    0.354
## Year2011          -0.01850    0.15630   -0.12    0.906
## Year2012          -0.02383    0.15606   -0.15    0.879
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.46
## Multiple R-squared:  0.0577, Adjusted R-squared:  0.0169
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 383 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.216  0.865   0.951   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      2.44e-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 1.157 1          1.076
## Year            1.157 16          1.005

```

Residuals from last author



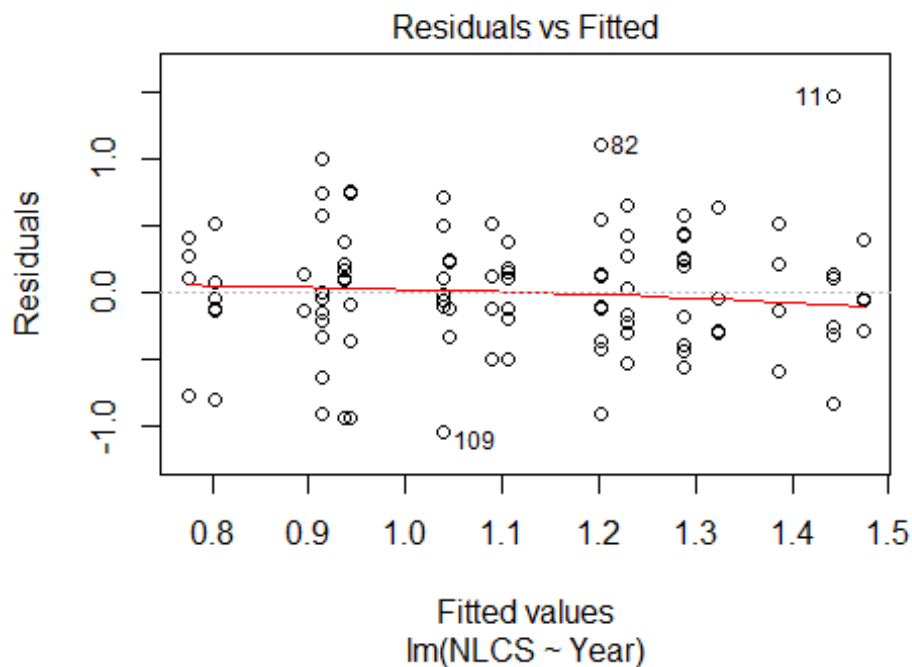
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3856 -0.3112 0.0359 0.3200 1.5503
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.260938 0.149061 8.46 5.4e-16 ***
## LastAuthorFemale1 0.001473 0.065310 0.02 0.982
## Year1997 0.219908 0.190879 1.15 0.250
## Year1998 0.034449 0.211927 0.16 0.871
## Year1999 0.067098 0.207900 0.32 0.747
## Year2000 -0.099374 0.271659 -0.37 0.715
## Year2001 -0.032348 0.209202 -0.15 0.877
## Year2002 -0.082901 0.173648 -0.48 0.633
## Year2003 -0.363282 0.193540 -1.88 0.061 .
## Year2004 0.123183 0.192189 0.64 0.522
## Year2005 -0.072673 0.185244 -0.39 0.695
## Year2006 0.001790 0.174135 0.01 0.992
```

```

## Year2007      0.017228    0.167170    0.10    0.918
## Year2008     -0.096088    0.164652   -0.58    0.560
## Year2009     -0.084721    0.173117   -0.49    0.625
## Year2010     -0.166864    0.188531   -0.89    0.377
## Year2011     -0.000112    0.160370    0.00    0.999
## Year2012     -0.004901    0.160455   -0.03    0.976
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.46
## Multiple R-squared:  0.0521, Adjusted R-squared:  0.0109
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 381 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.233  0.860  0.950  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      2.44e-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: 410"
## [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
##   14   21   11    8    8   16   12   12   13   11   13   15    8    7    9
## 2011 2012
##    6   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    2    4    4    4    9    8   10    6    7   11    4    4    7    6
## 2011 2012

```

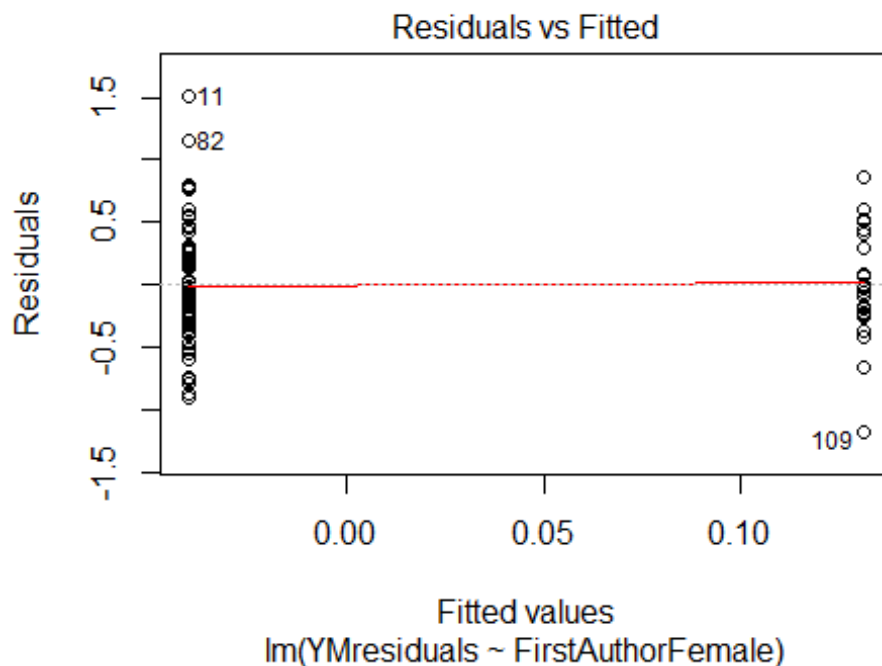
```
##      4      9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6      2      3      4      2      9      6      8      5      6      9      2      4      7      3
## 2011 2012
##      4      9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 11, df = 16, p-value = 0.8
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.17, df = 1, p-value = 0.7
## [1] "Female first author team size 2018 geometric mean: 2.46621207433047"
## [1] "Male first author team size 2018 geometric mean: 3.10369114783072"
## 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 = 7, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.0800838230519"
## [1] "Male last author team size 2018 geometric mean: 3.43754385517496"

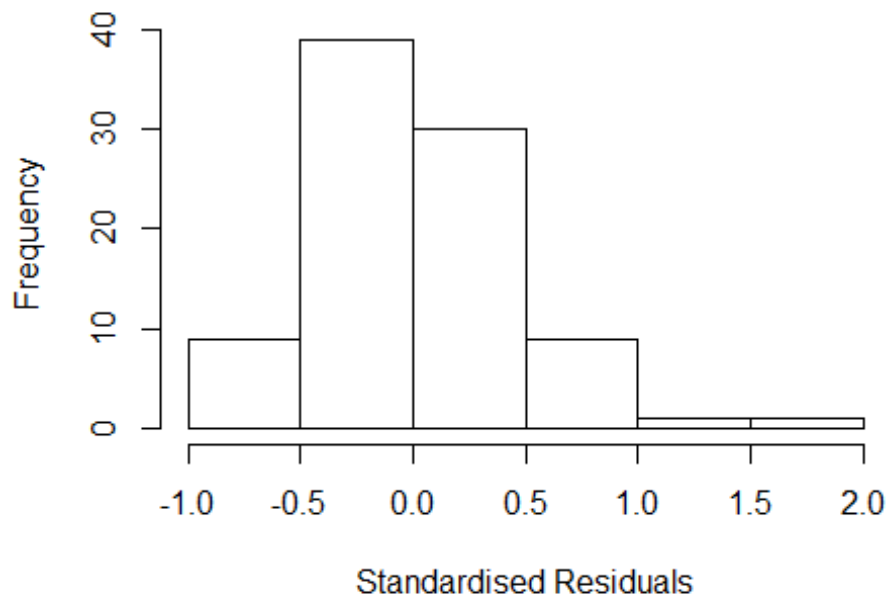
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, 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	2.852	1	1.689
LastAuthorFemale	1.860	1	1.364
UniqueAuthors	12.417	4	1.370
Year	27.144	16	1.109

Residuals from first and last author and team size



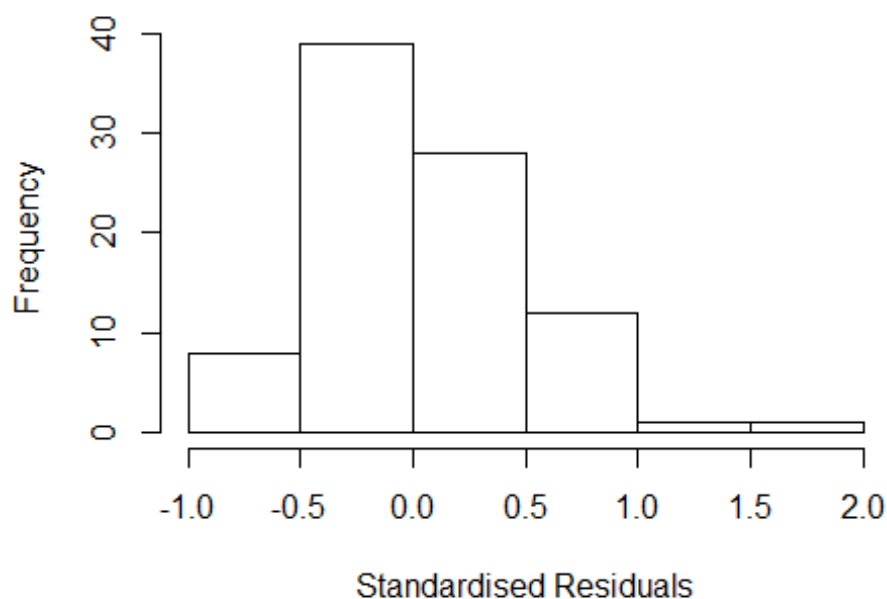
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.8131 -0.2407 -0.0433  0.2395  1.8814
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.027552   0.144877    7.09 1.1e-09 ***
## FirstAuthorFemale1 0.200407   0.121650    1.65  0.1042
## LastAuthorFemale1 0.009369   0.118916    0.08  0.9374
## UniqueAuthors2    0.147771   0.104976    1.41  0.1639
## UniqueAuthors3    0.182336   0.139033    1.31  0.1942
## UniqueAuthors4    0.647505   0.233492    2.77  0.0072 **
## UniqueAuthors5    0.482995   0.214543    2.25  0.0277 *
## Year1997         -0.223720   0.217986   -1.03  0.3085
## Year1998          0.168977   0.222697    0.76  0.4507
## Year1999         -0.126869   0.221931   -0.57  0.5695
```

```

## Year2000      0.232894  0.177217  1.31  0.1933
## Year2001     -0.061202  0.218308 -0.28  0.7801
## Year2002     -0.000518  0.178571  0.00  0.9977
## Year2003     -0.420732  0.183464 -2.29  0.0250 *
## Year2004     -0.335460  0.302424 -1.11  0.2714
## Year2005     -0.379723  0.214657 -1.77  0.0815 .
## Year2006     -0.043015  0.182369 -0.24  0.8143
## Year2007      0.003473  0.145696  0.02  0.9811
## Year2008     -0.455243  0.322829 -1.41  0.1632
## Year2009     -0.049611  0.177110 -0.28  0.7803
## Year2010     -0.286849  0.381094 -0.75  0.4543
## Year2011      0.209721  0.240287  0.87  0.3859
## Year2012     -0.169279  0.174940 -0.97  0.3368
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.472
## Multiple R-squared:  0.313, Adjusted R-squared:  0.0834
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.077  0.910  0.974  0.927  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      1.12e-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 2.378 1          1.542
## LastAuthorFemale  1.691 1          1.300
## Year              4.015 16          1.044

```

Residuals from first and last author



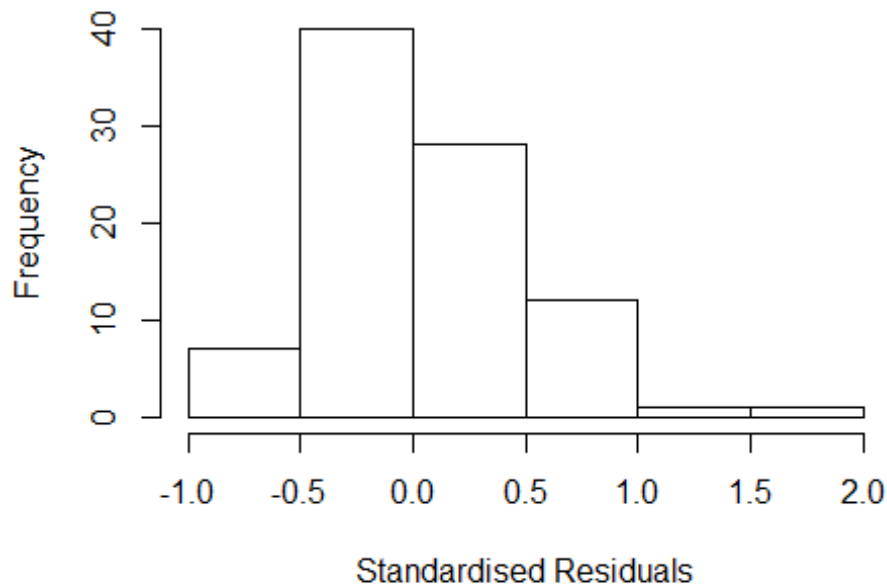
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8278 -0.2268 -0.0431 0.2955 1.7978
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.11121 0.13407 8.29 5.4e-12 ***
## FirstAuthorFemale1 0.27431 0.13311 2.06 0.043 *
## LastAuthorFemale1 0.04079 0.13762 0.30 0.768
## Year1997 -0.21621 0.16379 -1.32 0.191
## Year1998 0.22161 0.20001 1.11 0.272
## Year1999 -0.10514 0.23109 -0.45 0.651
## Year2000 0.31429 0.18566 1.69 0.095 .
## Year2001 -0.07367 0.21770 -0.34 0.736
## Year2002 -0.00604 0.17042 -0.04 0.972
## Year2003 -0.42515 0.22165 -1.92 0.059 .
## Year2004 -0.29165 0.44421 -0.66 0.514
## Year2005 -0.36789 0.23999 -1.53 0.130
```

```

## Year2006          0.08124    0.18343    0.44    0.659
## Year2007          0.03064    0.17704    0.17    0.863
## Year2008         -0.28344    0.29514   -0.96    0.340
## Year2009         -0.05000    0.17646   -0.28    0.778
## Year2010         -0.36576    0.47679   -0.77    0.446
## Year2011          0.21736    0.23287    0.93    0.354
## Year2012         -0.04389    0.18182   -0.24    0.810
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.239, Adjusted R-squared:  0.0433
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0742 0.8830 0.9610 0.9120 0.9880 0.9980
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.12e-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 2.419 1      1.555
## Year              2.419 16      1.028

```

Residuals from first author

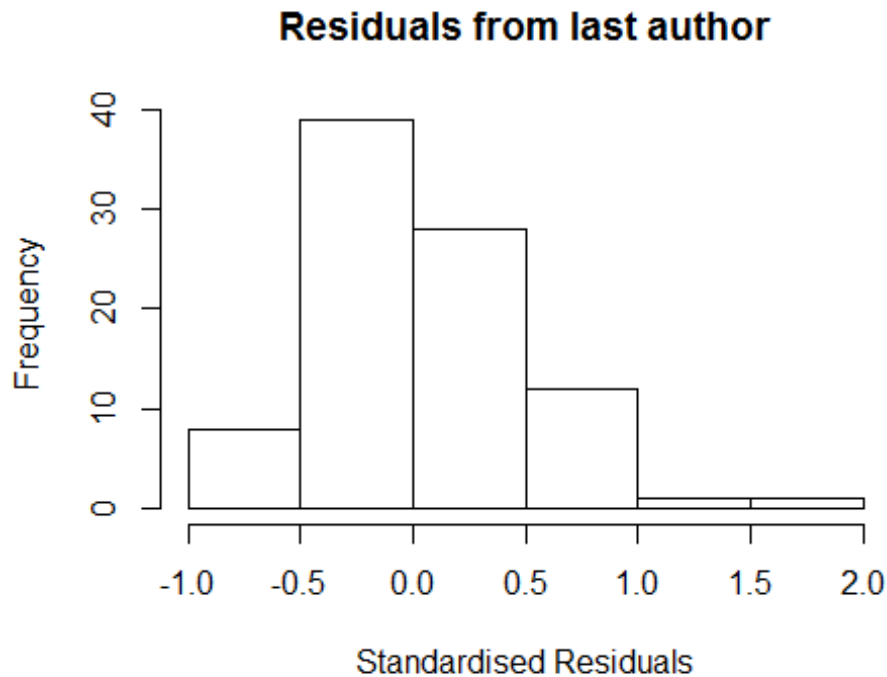


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.830 -0.233 -0.037 0.293 1.804
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10539 0.13025 8.49 2.1e-12 ***
## FirstAuthorFemale1 0.27963 0.13556 2.06 0.043 *
## Year1997 -0.21039 0.16070 -1.31 0.195
## Year1998 0.22571 0.19680 1.15 0.255
## Year1999 -0.10137 0.22896 -0.44 0.659
## Year2000 0.32011 0.18299 1.75 0.085 .
## Year2001 -0.06705 0.21639 -0.31 0.758
## Year2002 0.00662 0.16260 0.04 0.968
## Year2003 -0.41328 0.21313 -1.94 0.056 .
## Year2004 -0.27613 0.41031 -0.67 0.503
## Year2005 -0.35735 0.24402 -1.46 0.147
## Year2006 0.09177 0.18181 0.50 0.615
```

```

## Year2007          0.03380    0.17442    0.19    0.847
## Year2008          -0.27581    0.29234   -0.94    0.349
## Year2009          -0.03878    0.16998   -0.23    0.820
## Year2010          -0.33996    0.45238   -0.75    0.455
## Year2011           0.22018    0.23175    0.95    0.345
## Year2012          -0.02830    0.17216   -0.16    0.870
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.238, Adjusted R-squared:  0.0561
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0611 0.8830 0.9610 0.9100 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.12e-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 2.315 1          1.521
## Year              2.315 16          1.027

```



```
## [1] "List of 0 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
## -0.950 -0.212 -0.036  0.275  1.744
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1650     0.1817   6.41 1.4e-08 ***
## LastAuthorFemale1  0.0893     0.1432   0.62  0.535
## Year1997        -0.2700     0.2046  -1.32  0.191
## Year1998         0.2444     0.2996   0.82  0.417
## Year1999        -0.0767     0.2690  -0.29  0.776
## Year2000         0.2605     0.2226   1.17  0.246
## Year2001        -0.0155     0.2487  -0.06  0.950
## Year2002        -0.0696     0.2099  -0.33  0.741
## Year2003        -0.5033     0.2439  -2.06  0.043 *
## Year2004        -0.2154     0.5383  -0.40  0.690
## Year2005        -0.3698     0.2594  -1.43  0.158
## Year2006         0.0212     0.2195   0.10  0.923
```

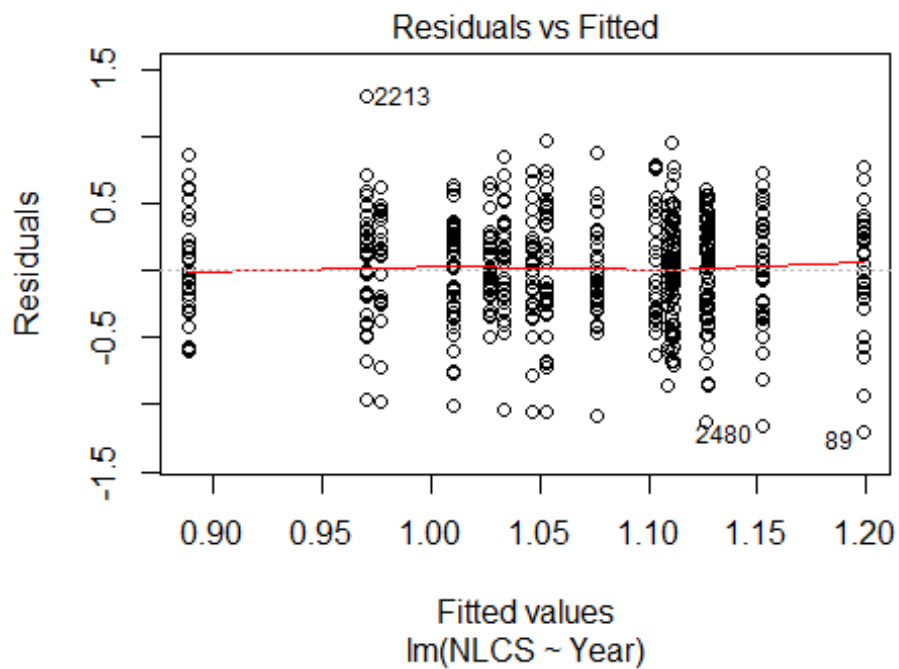
```

## Year2007          0.1140      0.1818      0.63      0.533
## Year2008          -0.3338      0.3229     -1.03      0.305
## Year2009          -0.0683      0.2191     -0.31      0.756
## Year2010          -0.2960      0.5166     -0.57      0.568
## Year2011           0.3044      0.2149      1.42      0.161
## Year2012           0.0227      0.2254      0.10      0.920
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.202, Adjusted R-squared:  0.0111
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 79 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0779 0.8820 0.9570 0.9020 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.12e-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: 89"
## [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
## 155 178 150 153 133 116 122 112 112 124 111 140 162 155 125
## 2011 2012
## 130 100
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 28 42 29 31 30 26 42 33 24 32 33 30 35 39 27
## 2011 2012

```



```
## 33 39
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 26 37 25 27 24 22 34 26 22 25 31 29 31 36 26
## 2011 2012
## 30 34
## [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
```



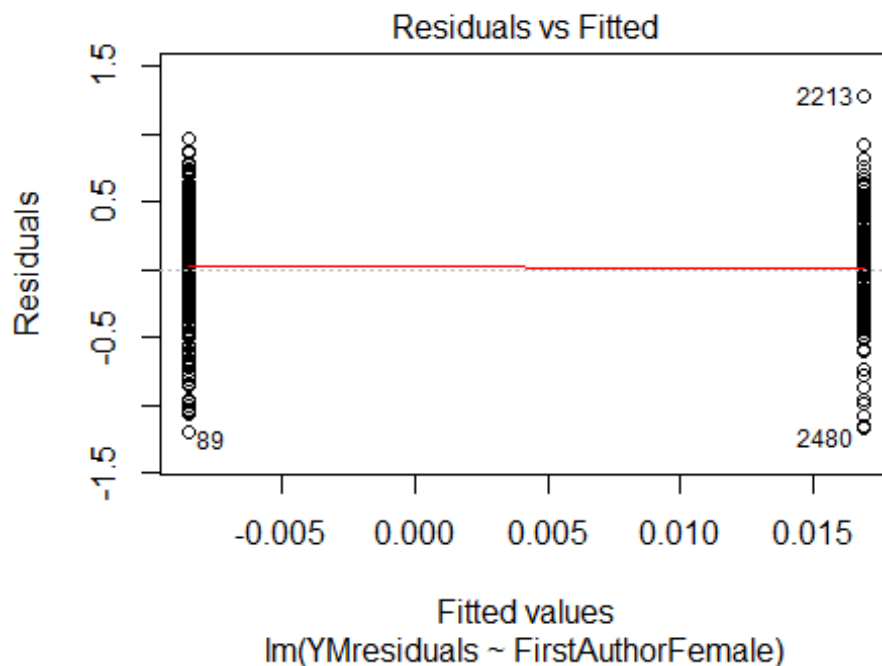
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.5, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 4.37482361574226"
## [1] "Male first author team size 2018 geometric mean: 4.19693866056616"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

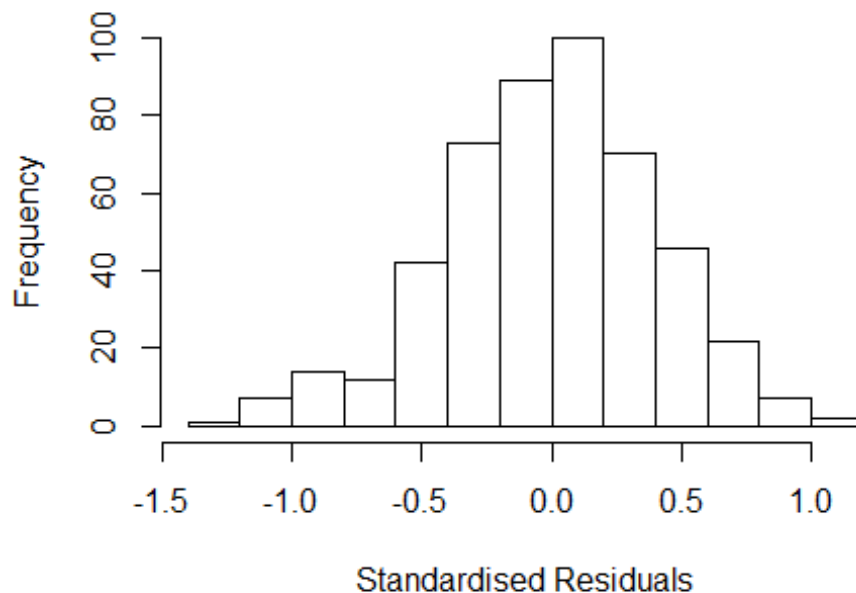
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 720, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.24502246292417"
## [1] "Male last author team size 2018 geometric mean: 4.27141771349555"

## 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 = 700, 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 1.144 1      1.070
## LastAuthorFemale  1.298 1      1.140
## UniqueAuthors    1.815 4      1.077
## Year              2.099 16     1.023
```

Residuals from first and last author and team size



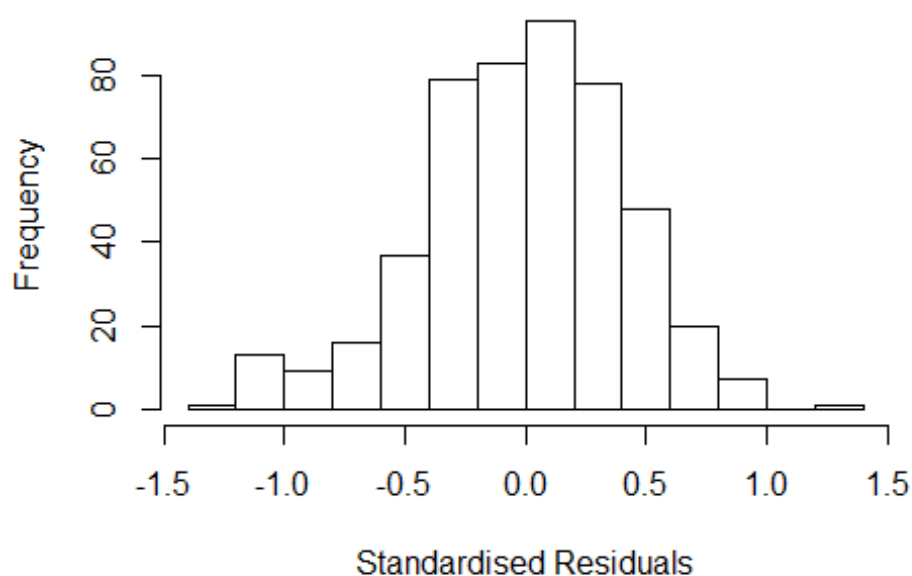
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.25137 -0.25784  0.00991  0.26114  1.16752
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.05750    0.11565   9.14  <2e-16 ***
## FirstAuthorFemale1  0.00399    0.03956   0.10  0.9196
## LastAuthorFemale1 -0.00226    0.04355  -0.05  0.9586
## UniqueAuthors2     0.04504    0.08854   0.51  0.6112
## UniqueAuthors3     0.16624    0.08423   1.97  0.0490 *
## UniqueAuthors4     0.26436    0.08521   3.10  0.0020 **
## UniqueAuthors5     0.26312    0.08785   3.00  0.0029 **
## Year1997         -0.06441    0.10487  -0.61  0.5394
## Year1998         -0.15010    0.12251  -1.23  0.2211
## Year1999         -0.11954    0.12064  -0.99  0.3222
```

```

## Year2000      -0.17697    0.10678   -1.66   0.0981 .
## Year2001      -0.11803    0.11836   -1.00   0.3192
## Year2002      -0.23168    0.11319   -2.05   0.0412 *
## Year2003      -0.18381    0.10670   -1.72   0.0856 .
## Year2004      -0.21275    0.13325   -1.60   0.1110
## Year2005      -0.17513    0.12075   -1.45   0.1477
## Year2006      -0.22281    0.12389   -1.80   0.0728 .
## Year2007      -0.07177    0.11592   -0.62   0.5361
## Year2008      -0.33948    0.11166   -3.04   0.0025 **
## Year2009      -0.22413    0.11848   -1.89   0.0591 .
## Year2010      -0.15985    0.13537   -1.18   0.2383
## Year2011      -0.07098    0.12410   -0.57   0.5676
## Year2012      -0.17067    0.13529   -1.26   0.2078
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.378
## Multiple R-squared:  0.0942, Adjusted R-squared:  0.0511
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 47 weights are ~= 1. The remaining 438 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.252  0.869  0.947  0.893  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      2.06e-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 1.142 1      1.069
## LastAuthorFemale  1.149 1      1.072
## Year              1.300 16      1.008

```

Residuals from first and last author



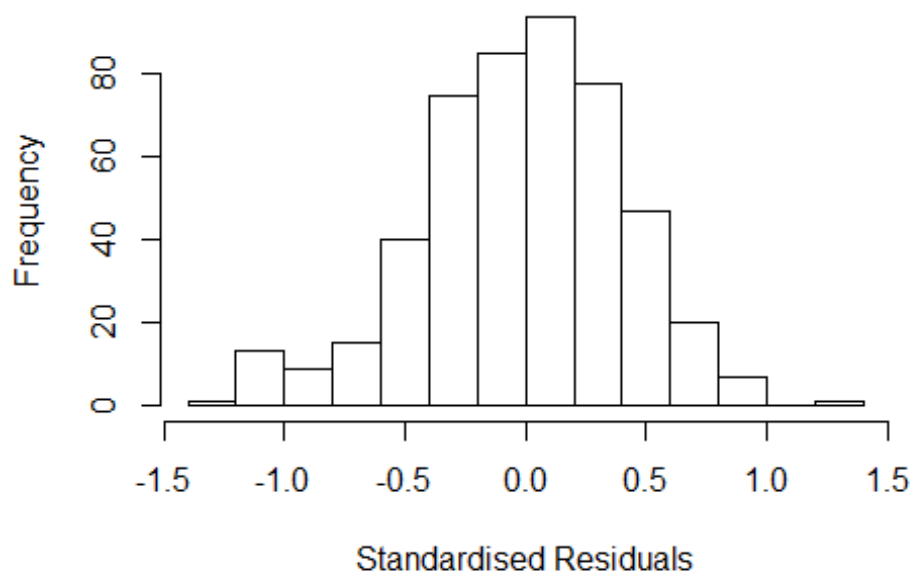
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22090 -0.27028 0.00941 0.27841 1.25899
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2209 0.1062 11.50 <2e-16 ***
## FirstAuthorFemale1 0.0100 0.0410 0.24 0.807
## LastAuthorFemale1 -0.0168 0.0430 -0.39 0.696
## Year1997 -0.0648 0.1206 -0.54 0.591
## Year1998 -0.1504 0.1381 -1.09 0.277
## Year1999 -0.0968 0.1326 -0.73 0.466
## Year2000 -0.1803 0.1173 -1.54 0.125
## Year2001 -0.1304 0.1298 -1.00 0.316
## Year2002 -0.2280 0.1261 -1.81 0.071 .
## Year2003 -0.1789 0.1237 -1.45 0.149
## Year2004 -0.1595 0.1432 -1.11 0.266
## Year2005 -0.1919 0.1330 -1.44 0.150
```

```

## Year2006          -0.1759      0.1340   -1.31    0.190
## Year2007          -0.0563      0.1301   -0.43    0.665
## Year2008          -0.3229      0.1292   -2.50    0.013 *
## Year2009          -0.2219      0.1326   -1.67    0.095 .
## Year2010          -0.1398      0.1455   -0.96    0.337
## Year2011          -0.0485      0.1397   -0.35    0.728
## Year2012          -0.1407      0.1498   -0.94    0.348
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0348, Adjusted R-squared:  -0.00245
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 451 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.307  0.875   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      2.06e-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 1.143 1      1.069
## Year              1.143 16      1.004

```

Residuals from first author



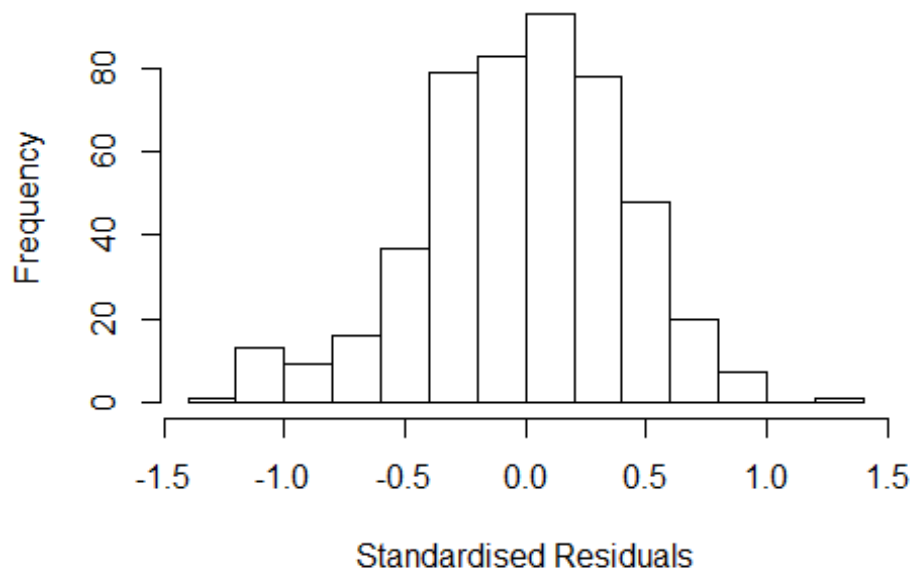
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2192 -0.2773  0.0111  0.2792  1.2628
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21917    0.10610   11.49  <2e-16 ***
## FirstAuthorFemale1 0.00836    0.04100    0.20   0.838
## Year1997      -0.06802    0.12063   -0.56   0.573
## Year1998      -0.15333    0.13801   -1.11   0.267
## Year1999      -0.09804    0.13261   -0.74   0.460
## Year2000      -0.18030    0.11740   -1.54   0.125
## Year2001      -0.13082    0.13021   -1.00   0.316
## Year2002      -0.22991    0.12595   -1.83   0.069 .
## Year2003      -0.18279    0.12300   -1.49   0.138
## Year2004      -0.16068    0.14350   -1.12   0.263
## Year2005      -0.19272    0.13267   -1.45   0.147
## Year2006      -0.17592    0.13408   -1.31   0.190
```

```

## Year2007          -0.05886    0.13018   -0.45    0.651
## Year2008          -0.32685    0.12839   -2.55    0.011 *
## Year2009          -0.22233    0.13251   -1.68    0.094 .
## Year2010          -0.14470    0.14431   -1.00    0.317
## Year2011          -0.05219    0.13935   -0.37    0.708
## Year2012          -0.14299    0.15033   -0.95    0.342
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0347, Adjusted R-squared:  -0.00044
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 35 weights are ~= 1. The remaining 450 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.303  0.874  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      2.06e-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 1.147 1          1.071
## Year            1.147 16          1.004

```


Residuals from last author



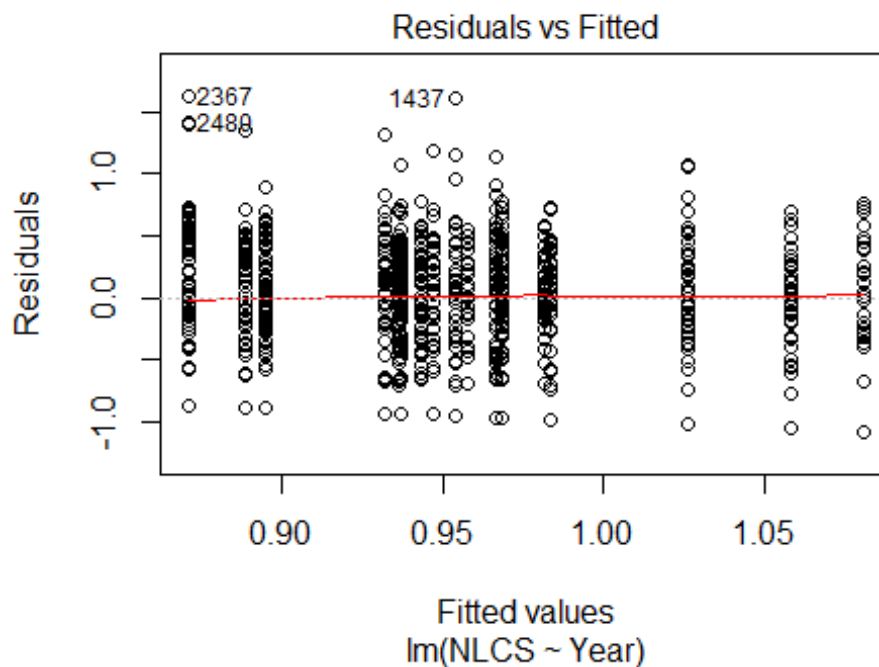
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.223 -0.271 0.011 0.277 1.266
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2233 0.1063 11.51 <2e-16 ***
## LastAuthorFemale1 -0.0156 0.0430 -0.36 0.717
## Year1997 -0.0637 0.1204 -0.53 0.597
## Year1998 -0.1505 0.1383 -1.09 0.277
## Year1999 -0.0966 0.1328 -0.73 0.467
## Year2000 -0.1815 0.1176 -1.54 0.123
## Year2001 -0.1308 0.1302 -1.00 0.316
## Year2002 -0.2261 0.1254 -1.80 0.072 .
## Year2003 -0.1800 0.1240 -1.45 0.147
## Year2004 -0.1562 0.1408 -1.11 0.268
## Year2005 -0.1908 0.1330 -1.43 0.152
## Year2006 -0.1745 0.1337 -1.30 0.193
```

```

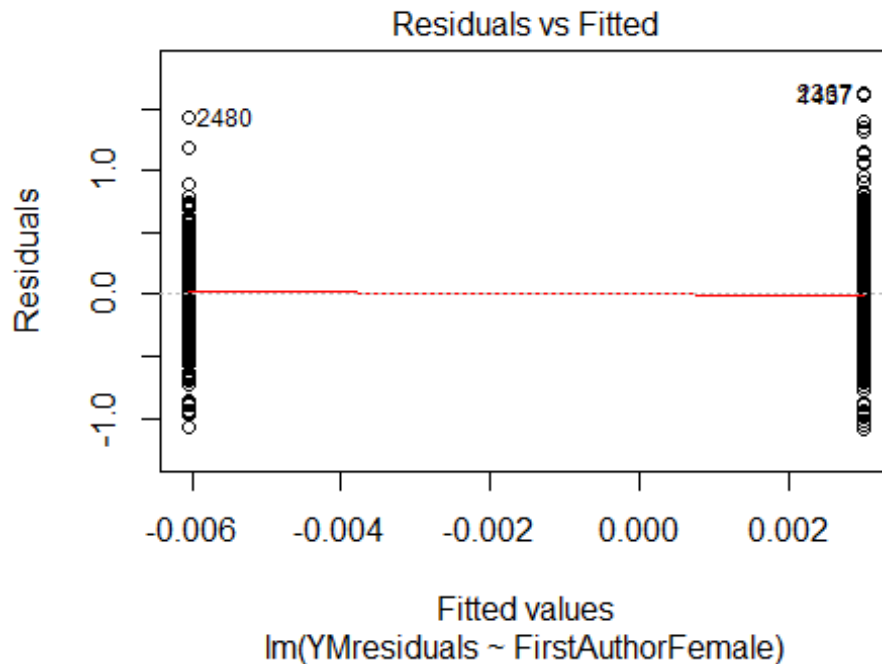
## Year2007          -0.0554      0.1300   -0.43    0.670
## Year2008          -0.3221      0.1289   -2.50    0.013 *
## Year2009          -0.2210      0.1324   -1.67    0.096 .
## Year2010          -0.1391      0.1451   -0.96    0.338
## Year2011          -0.0478      0.1395   -0.34    0.732
## Year2012          -0.1405      0.1500   -0.94    0.350
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0347, Adjusted R-squared:  -0.000411
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 451 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.300  0.873   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      2.06e-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: 485"
## [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
## 148 187 185 180 195 131 152 114 129 156 158 161 171 162 198
## 2011 2012
## 172 143
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 33 48 50 49 43 25 63 39 45 55 70 65 78 70 86
## 2011 2012

```

```
## 60 49
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 32 43 43 44 39 22 57 35 39 49 63 59 65 63 79
## 2011 2012
## 55 46
## [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
```



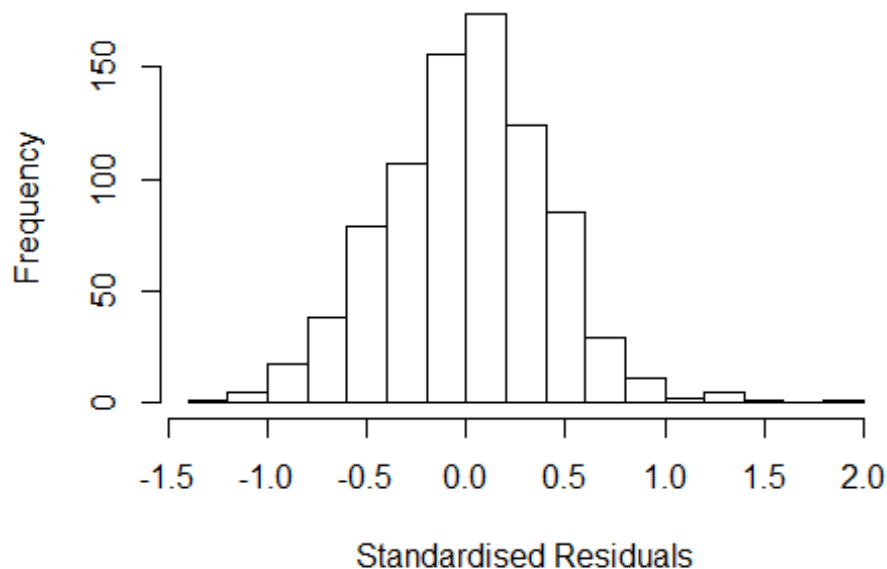
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.5, df = 1, p-value = 0.2
## [1] "Female first author team size 2018 geometric mean: 3.27272961778769"
## [1] "Male first author team size 2018 geometric mean: 3.14864313128207"
## 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 = 950, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.25654397963815"
## [1] "Male last author team size 2018 geometric mean: 3.17108201073556"
##
## 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] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF^(1/(2*Df))
FirstAuthorFemale	1.139	1	1.067
LastAuthorFemale	1.067	1	1.033
UniqueAuthors	1.425	4	1.045
Year	1.503	16	1.013

Residuals from first and last author and team size



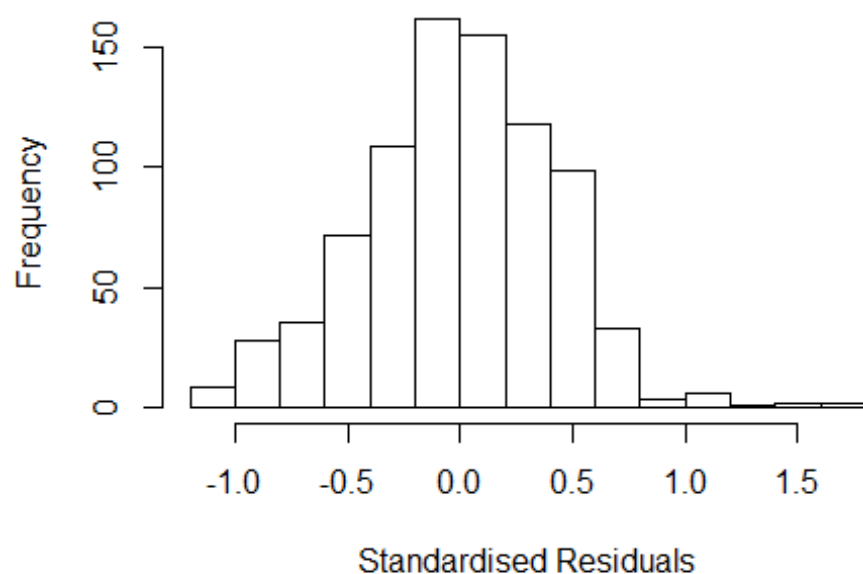
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2286 -0.2794 0.0133 0.2619 1.8050
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9769 0.0883 11.07 < 2e-16 ***
## FirstAuthorFemale1 -0.0433 0.0308 -1.41 0.1602
## LastAuthorFemale1 0.0335 0.0348 0.96 0.3360
## UniqueAuthors2 0.2037 0.0378 5.38 9.5e-08 ***
## UniqueAuthors3 0.2517 0.0447 5.63 2.5e-08 ***
## UniqueAuthors4 0.3066 0.0528 5.80 9.3e-09 ***
## UniqueAuthors5 0.3451 0.0500 6.90 1.1e-11 ***
## Year1997 -0.2118 0.1122 -1.89 0.0594 .
## Year1998 -0.0807 0.1099 -0.73 0.4628
## Year1999 -0.1121 0.0978 -1.15 0.2522
```

```

## Year2000          -0.0438      0.1070   -0.41   0.6825
## Year2001          -0.1666      0.1150   -1.45   0.1478
## Year2002          -0.1968      0.0983   -2.00   0.0456 *
## Year2003          -0.2111      0.1298   -1.63   0.1045
## Year2004          -0.1480      0.0999   -1.48   0.1389
## Year2005          -0.2568      0.1085   -2.37   0.0181 *
## Year2006          -0.2184      0.0995   -2.19   0.0285 *
## Year2007          -0.2164      0.1016   -2.13   0.0336 *
## Year2008          -0.2859      0.1085   -2.63   0.0086 **
## Year2009          -0.1433      0.1004   -1.43   0.1537
## Year2010          -0.2686      0.1000   -2.69   0.0074 **
## Year2011          -0.1506      0.1104   -1.36   0.1728
## Year2012          -0.2156      0.1110   -1.94   0.0525 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.108, Adjusted R-squared:  0.0836
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 75 weights are ~= 1. The remaining 758 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0068 0.8770 0.9490 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          1.20e-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 1.112 1 1.054
## LastAuthorFemale 1.059 1 1.029
## Year 1.165 16 1.005

```

Residuals from first and last author



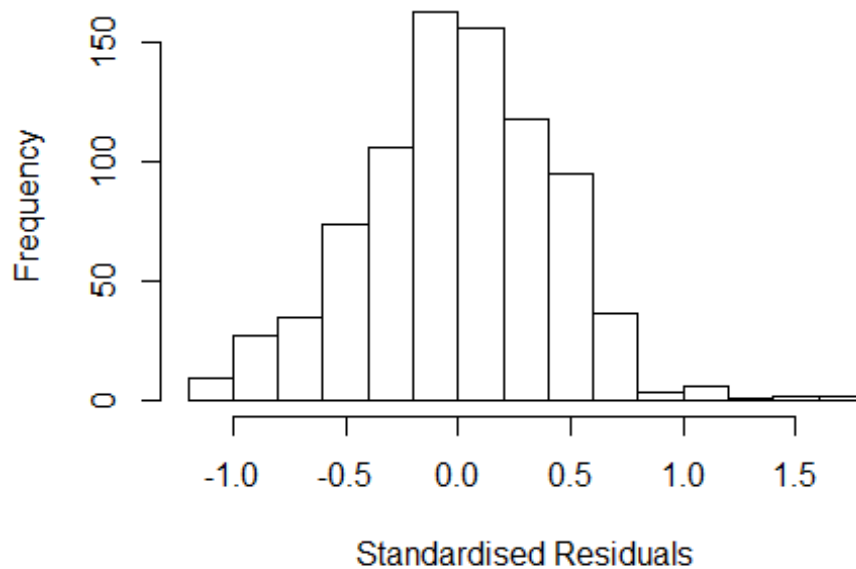
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.10941 -0.28036 0.00308 0.28522 1.67747
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08779 0.09257 11.75 <2e-16 ***
## FirstAuthorFemale1 0.00497 0.03164 0.16 0.875
## LastAuthorFemale1 0.01665 0.03750 0.44 0.657
## Year1997 -0.19535 0.11828 -1.65 0.099 .
## Year1998 -0.11047 0.11457 -0.96 0.335
## Year1999 -0.10697 0.10237 -1.05 0.296
## Year2000 0.00496 0.11183 0.04 0.965
## Year2001 -0.14087 0.11859 -1.19 0.235
## Year2002 -0.14252 0.10376 -1.37 0.170
## Year2003 -0.20525 0.13553 -1.51 0.130
## Year2004 -0.08543 0.10640 -0.80 0.422
## Year2005 -0.19863 0.11336 -1.75 0.080 .
```

```

## Year2006      -0.15141    0.10526   -1.44    0.151
## Year2007      -0.15642    0.10601   -1.48    0.140
## Year2008      -0.24349    0.11702   -2.08    0.038 *
## Year2009      -0.08866    0.10728   -0.83    0.409
## Year2010      -0.19384    0.10580   -1.83    0.067 .
## Year2011      -0.07787    0.11954   -0.65    0.515
## Year2012      -0.13323    0.11679   -1.14    0.254
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.023, Adjusted R-squared:  0.00138
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 80 weights are ~= 1. The remaining 753 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0656 0.8640 0.9470 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      1.20e-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 1.109 1      1.053
## Year      1.109 16      1.003

```


Residuals from first author



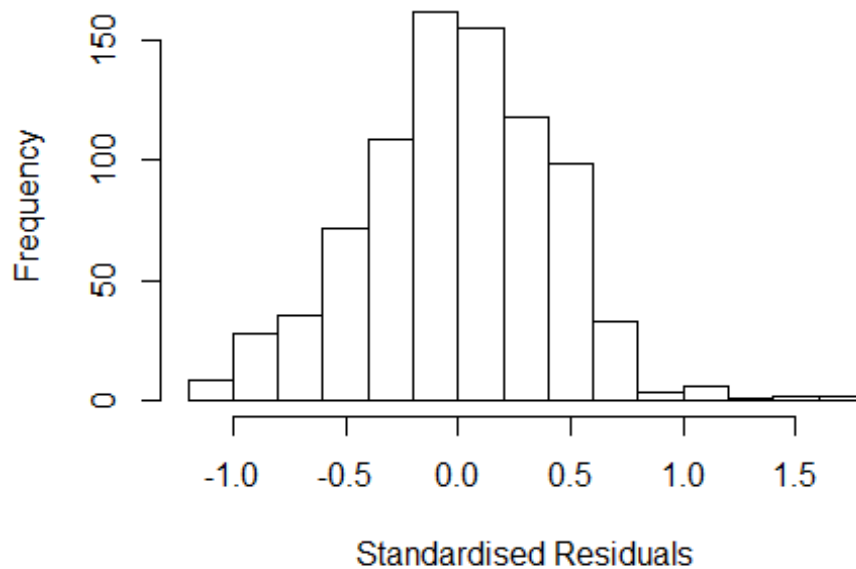
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.09762 -0.27922 0.00462 0.29238 1.67661
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08910 0.09246 11.78 <2e-16 ***
## FirstAuthorFemale1 0.00852 0.03200 0.27 0.790
## Year1997 -0.19448 0.11851 -1.64 0.101
## Year1998 -0.10891 0.11423 -0.95 0.341
## Year1999 -0.10531 0.10227 -1.03 0.303
## Year2000 0.00600 0.11182 0.05 0.957
## Year2001 -0.14072 0.11881 -1.18 0.237
## Year2002 -0.14257 0.10371 -1.37 0.170
## Year2003 -0.20571 0.13567 -1.52 0.130
## Year2004 -0.08487 0.10630 -0.80 0.425
## Year2005 -0.19553 0.11332 -1.73 0.085 .
## Year2006 -0.15180 0.10526 -1.44 0.150
```

```

## Year2007          -0.15448    0.10574   -1.46    0.144
## Year2008          -0.24325    0.11693   -2.08    0.038 *
## Year2009          -0.08711    0.10736   -0.81    0.417
## Year2010          -0.19352    0.10576   -1.83    0.068 .
## Year2011          -0.07783    0.11951   -0.65    0.515
## Year2012          -0.13150    0.11648   -1.13    0.259
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.0228, Adjusted R-squared:  0.0024
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 80 weights are ~= 1. The remaining 753 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0658 0.8640 0.9460 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      1.20e-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 1.055 1          1.027
## Year            1.055 16          1.002

```

Residuals from last author



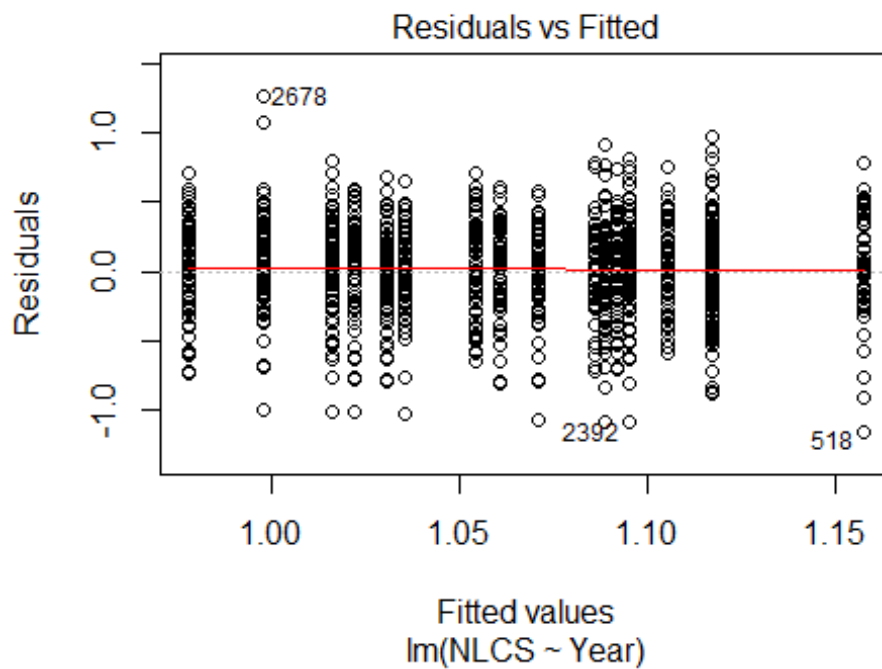
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.10626 -0.28152 0.00333 0.28331 1.67577
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08818 0.09256 11.76 <2e-16 ***
## LastAuthorFemale1 0.01808 0.03786 0.48 0.633
## Year1997 -0.19551 0.11821 -1.65 0.099 .
## Year1998 -0.10999 0.11424 -0.96 0.336
## Year1999 -0.10632 0.10212 -1.04 0.298
## Year2000 0.00568 0.11143 0.05 0.959
## Year2001 -0.14010 0.11804 -1.19 0.236
## Year2002 -0.14089 0.10264 -1.37 0.170
## Year2003 -0.20396 0.13446 -1.52 0.130
## Year2004 -0.08467 0.10583 -0.80 0.424
## Year2005 -0.19808 0.11300 -1.75 0.080 .
## Year2006 -0.15042 0.10452 -1.44 0.150
```

```

## Year2007          -0.15534      0.10530    -1.48      0.141
## Year2008          -0.24209      0.11616    -2.08      0.037 *
## Year2009          -0.08756      0.10655    -0.82      0.411
## Year2010          -0.19249      0.10487    -1.84      0.067 .
## Year2011          -0.07597      0.11767    -0.65      0.519
## Year2012          -0.13258      0.11652    -1.14      0.256
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.023, Adjusted R-squared:  0.00258
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 81 weights are ~= 1. The remaining 752 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0662 0.8640 0.9460 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      1.20e-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: 833"
## [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
## 115 129 109 125 131 178 149 156 101 122 130 98 119 120 113
## 2011 2012
## 135 161
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 64 62 66 66 58 78 82 97 73 81 87 61 73 82 74
## 2011 2012

```

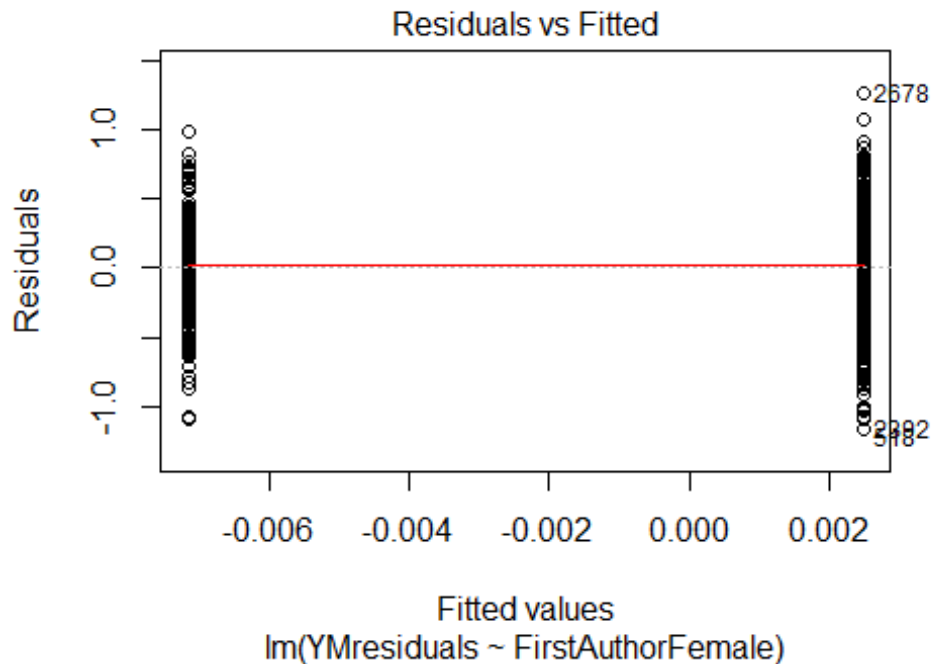
```
## 79 92
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 55 58 61 57 51 75 74 90 71 71 76 54 66 78 67
## 2011 2012
## 70 89
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.5
```



```
##
## 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: 3.11337110469392"
## [1] "Male first author team size 2018 geometric mean: 3.12609230125565"

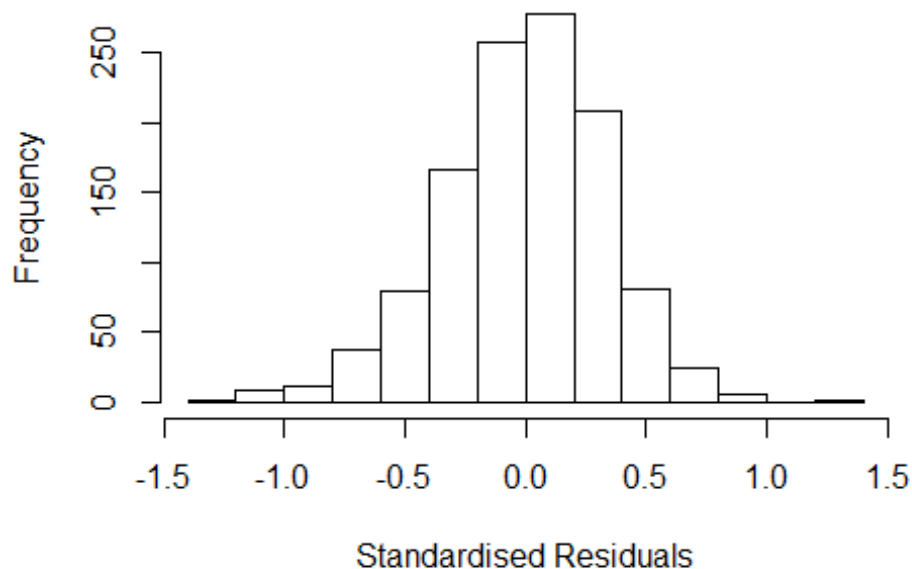
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 590, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.37448783970553"
## [1] "Male last author team size 2018 geometric mean: 3.36686937710805"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 280, 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	1.209	1	1.099
LastAuthorFemale	1.148	1	1.072
UniqueAuthors	1.503	4	1.052
Year	1.667	16	1.016

Residuals from first and last author and team size



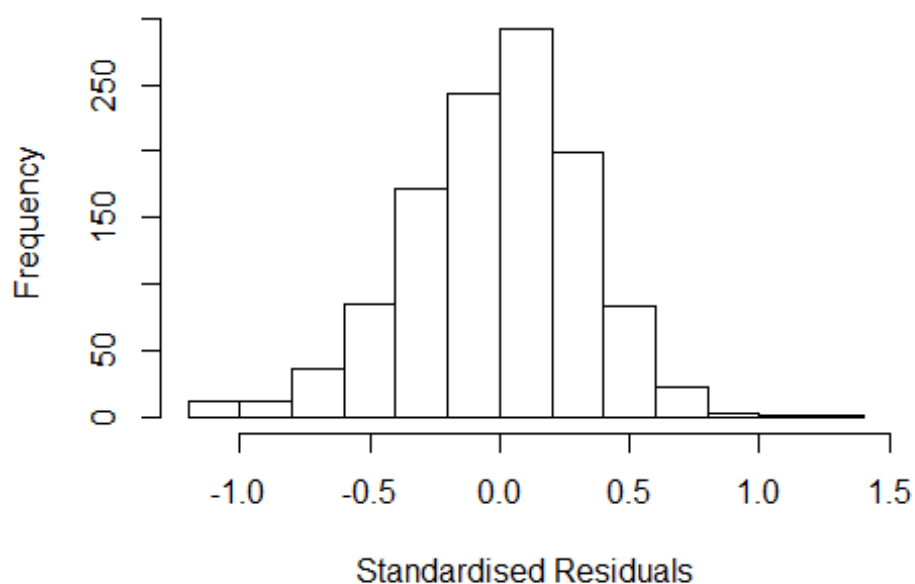
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2287 -0.2147  0.0137  0.2190  1.2740
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.01166    0.05497   18.40 < 2e-16 ***
## FirstAuthorFemale1 -0.02007    0.02341   -0.86  0.39139
## LastAuthorFemale1 -0.00224    0.02504   -0.09  0.92887
## UniqueAuthors2     0.10624    0.02988    3.56  0.00039 ***
## UniqueAuthors3     0.09127    0.03032    3.01  0.00267 **
## UniqueAuthors4     0.10695    0.03629    2.95  0.00327 **
## UniqueAuthors5     0.20105    0.04109    4.89  1.1e-06 ***
## Year1997          0.04833    0.06918    0.70  0.48492
## Year1998          0.01345    0.06846    0.20  0.84432
## Year1999          0.11081    0.06924    1.60  0.10979
```

```

## Year2000          0.02100      0.07488      0.28  0.77917
## Year2001         -0.04837      0.06662     -0.73  0.46788
## Year2002         -0.03049      0.06589     -0.46  0.64363
## Year2003          0.01637      0.06023      0.27  0.78588
## Year2004         -0.03199      0.06626     -0.48  0.62937
## Year2005         -0.08055      0.06664     -1.21  0.22699
## Year2006         -0.04322      0.06476     -0.67  0.50465
## Year2007         -0.03439      0.06847     -0.50  0.61559
## Year2008         -0.04769      0.06334     -0.75  0.45167
## Year2009          0.02131      0.06169      0.35  0.72979
## Year2010          0.01359      0.06267      0.22  0.82838
## Year2011          0.01533      0.06845      0.22  0.82285
## Year2012         -0.11169      0.06389     -1.75  0.08069 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.323
## Multiple R-squared:  0.0428, Adjusted R-squared:  0.0243
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 98 weights are ~= 1. The remaining 1065 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0836 0.8730 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      8.60e-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 1.162 1      1.078
## LastAuthorFemale  1.128 1      1.062
## Year              1.193 16      1.006

```


Residuals from first and last author



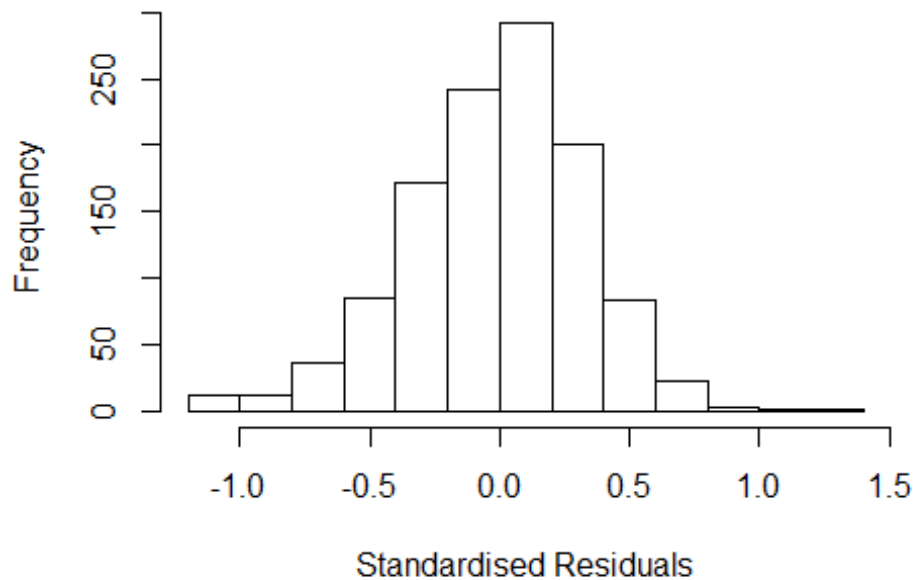
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1932 -0.2247 0.0145 0.2197 1.2469
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0805 0.0500 21.62 <2e-16 ***
## FirstAuthorFemale1 -0.0112 0.0231 -0.49 0.63
## LastAuthorFemale1 0.0049 0.0251 0.20 0.85
## Year1997 0.0488 0.0688 0.71 0.48
## Year1998 0.0177 0.0674 0.26 0.79
## Year1999 0.1127 0.0707 1.59 0.11
## Year2000 0.0222 0.0738 0.30 0.76
## Year2001 -0.0424 0.0654 -0.65 0.52
## Year2002 -0.0251 0.0658 -0.38 0.70
## Year2003 0.0247 0.0597 0.41 0.68
## Year2004 -0.0120 0.0655 -0.18 0.86
## Year2005 -0.0589 0.0647 -0.91 0.36
```

```

## Year2006          -0.0273      0.0622   -0.44      0.66
## Year2007          -0.0109      0.0669   -0.16      0.87
## Year2008          -0.0269      0.0615   -0.44      0.66
## Year2009           0.0297      0.0596     0.50      0.62
## Year2010           0.0343      0.0615     0.56      0.58
## Year2011           0.0478      0.0650     0.74      0.46
## Year2012          -0.0693      0.0624   -1.11      0.27
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.327
## Multiple R-squared:  0.0176, Adjusted R-squared:  0.00211
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 100 weights are ~= 1. The remaining 1063 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
## 0.115 0.873 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      8.60e-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 1.102 1      1.050
## Year              1.102 16      1.003

```

Residuals from first author



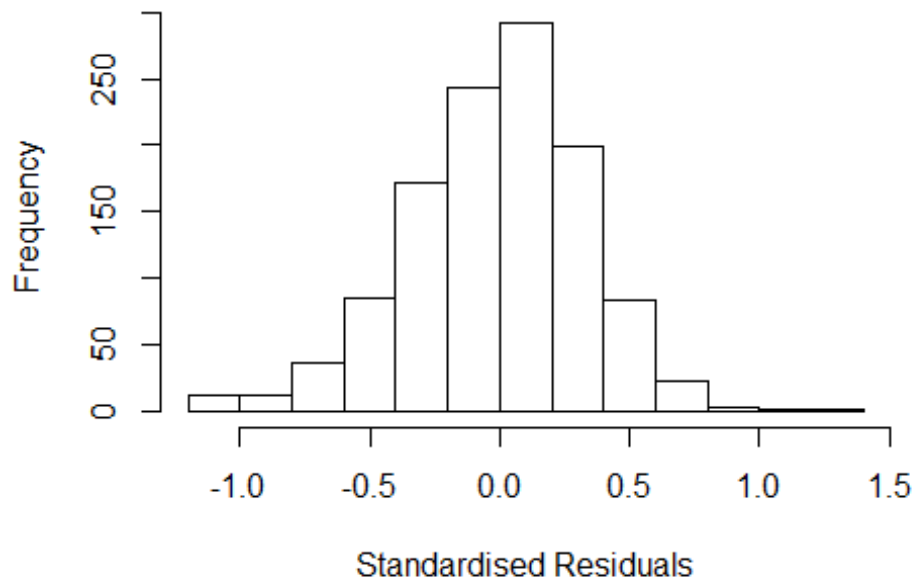
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1942 -0.2252 0.0141 0.2191 1.2510
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0812 0.0498 21.73 <2e-16 ***
## FirstAuthorFemale1 -0.0103 0.0225 -0.46 0.65
## Year1997 0.0495 0.0686 0.72 0.47
## Year1998 0.0173 0.0673 0.26 0.80
## Year1999 0.1130 0.0706 1.60 0.11
## Year2000 0.0227 0.0737 0.31 0.76
## Year2001 -0.0422 0.0654 -0.65 0.52
## Year2002 -0.0253 0.0658 -0.39 0.70
## Year2003 0.0248 0.0597 0.42 0.68
## Year2004 -0.0116 0.0655 -0.18 0.86
## Year2005 -0.0593 0.0646 -0.92 0.36
## Year2006 -0.0271 0.0621 -0.44 0.66
```

```

## Year2007          -0.0110      0.0669   -0.16      0.87
## Year2008          -0.0270      0.0615   -0.44      0.66
## Year2009           0.0298      0.0596    0.50      0.62
## Year2010           0.0340      0.0613    0.55      0.58
## Year2011           0.0482      0.0649    0.74      0.46
## Year2012          -0.0692      0.0623   -1.11      0.27
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.327
## Multiple R-squared:  0.0176, Adjusted R-squared:  0.00298
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 101 weights are ~= 1. The remaining 1062 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.112  0.872  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      8.60e-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 1.07 1          1.034
## Year              1.07 16          1.002

```

Residuals from last author



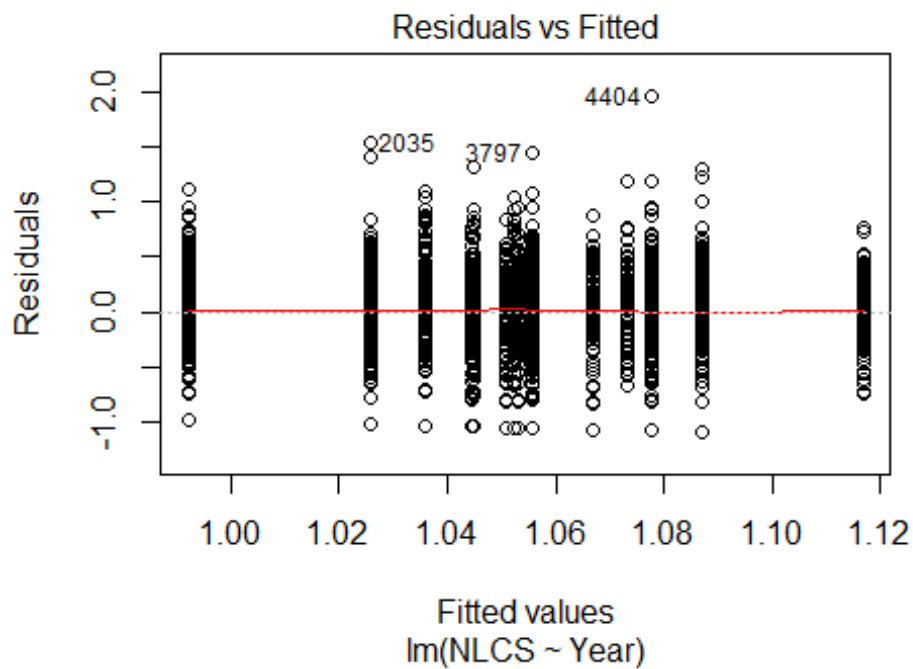
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1910 -0.2232 0.0155 0.2189 1.2530
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.07872 0.04999 21.58 <2e-16 ***
## LastAuthorFemale1 0.00239 0.02448 0.10 0.92
## Year1997 0.04992 0.06877 0.73 0.47
## Year1998 0.01755 0.06750 0.26 0.79
## Year1999 0.11225 0.07081 1.59 0.11
## Year2000 0.02374 0.07376 0.32 0.75
## Year2001 -0.04310 0.06556 -0.66 0.51
## Year2002 -0.02470 0.06584 -0.38 0.71
## Year2003 0.02493 0.05982 0.42 0.68
## Year2004 -0.01328 0.06527 -0.20 0.84
## Year2005 -0.05948 0.06471 -0.92 0.36
## Year2006 -0.02848 0.06222 -0.46 0.65
```

```

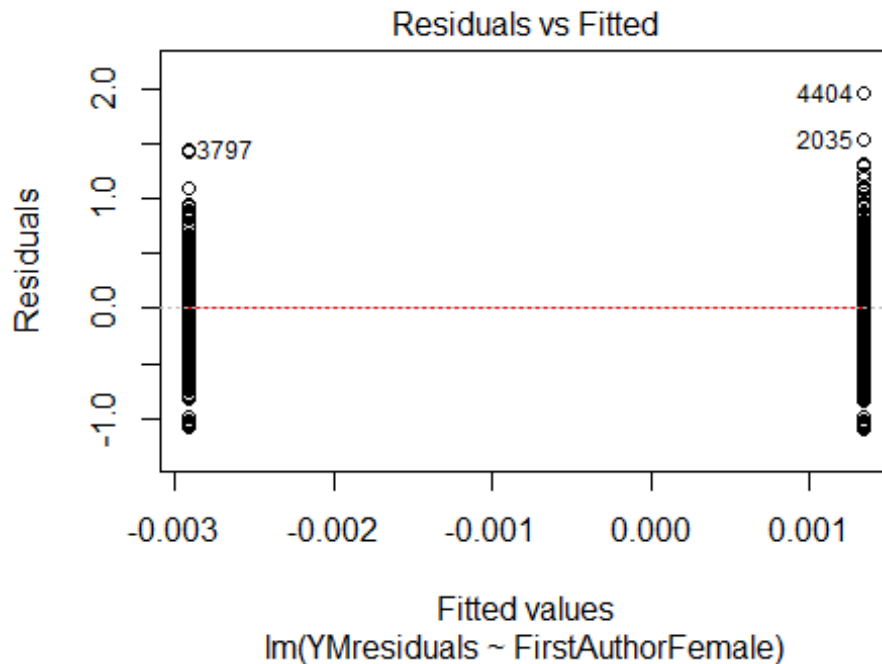
## Year2007          -0.01249      0.06689    -0.19      0.85
## Year2008          -0.02843      0.06142    -0.46      0.64
## Year2009           0.02915      0.05965     0.49      0.63
## Year2010           0.03149      0.06094     0.52      0.61
## Year2011           0.04617      0.06495     0.71      0.48
## Year2012          -0.07115      0.06214    -1.14      0.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.327
## Multiple R-squared:  0.0174, Adjusted R-squared:  0.00281
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 102 weights are ~= 1. The remaining 1061 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.110  0.872  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      8.60e-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: 1163"
## [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
##  214  221  223  244  238  255  244  246  189  222  216  204  256  243  284
## 2011 2012
##  273  286
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   85  100  108  129  109  117  148  163  116  149  156  141  169  154  193
## 2011 2012

```

```
## 176 180
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 76 91 102 115 98 108 134 150 110 134 139 124 153 135 176
## 2011 2012
## 167 168
## [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
```

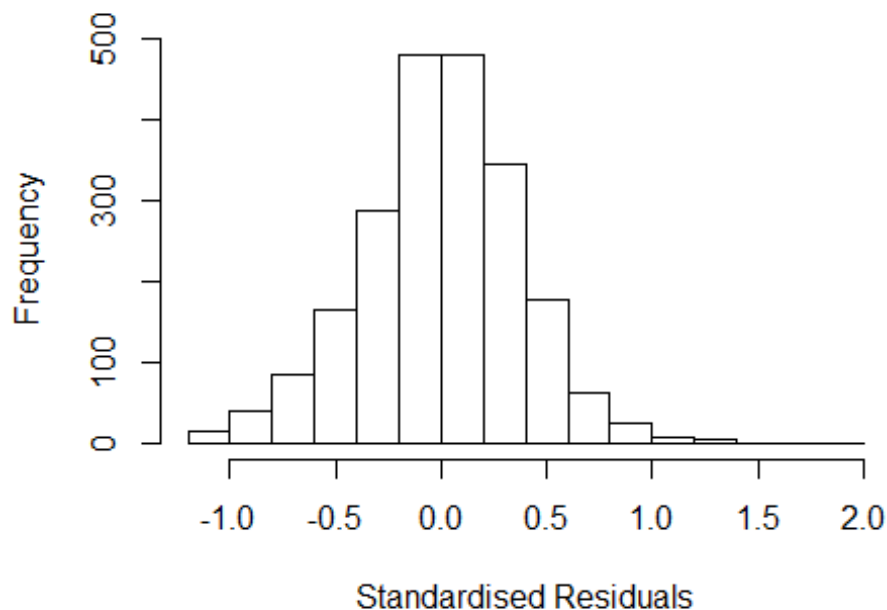


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.8, df = 1, p-value = 0.002
```



```
## [1] "Female first author team size 2018 geometric mean: 3.14749916336737"
## [1] "Male first author team size 2018 geometric mean: 3.12559986936115"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2400, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.06068007546025"
## [1] "Male last author team size 2018 geometric mean: 3.16730725277136"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2100, 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 1.123 1 1.060
## LastAuthorFemale 1.071 1 1.035
## UniqueAuthors 1.266 4 1.030
## Year 1.410 16 1.011
```


Residuals from first and last author and team size



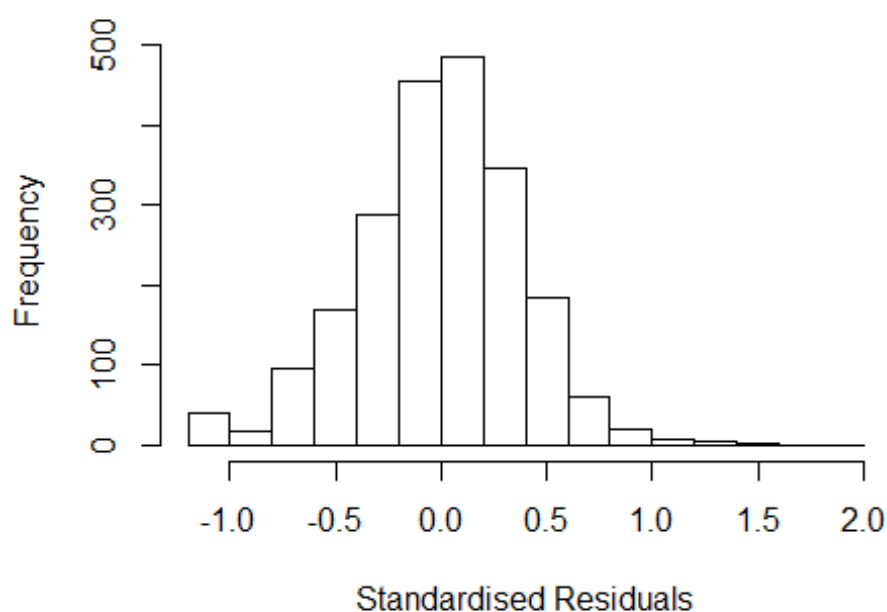
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.14938 -0.22856  0.00598  0.23970  1.95797
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.93384    0.04736   19.72  < 2e-16 ***
## FirstAuthorFemale1 -0.01321    0.01738   -0.76    0.45
## LastAuthorFemale1 -0.02377    0.01865   -1.27    0.20
## UniqueAuthors2     0.16022    0.02480    6.46 1.3e-10 ***
## UniqueAuthors3     0.15850    0.02611    6.07 1.5e-09 ***
## UniqueAuthors4     0.20574    0.03011    6.83 1.1e-11 ***
## UniqueAuthors5     0.25325    0.03433    7.38 2.3e-13 ***
## Year1997           0.01574    0.05961    0.26    0.79
## Year1998           0.01872    0.06118    0.31    0.76
## Year1999           0.04213    0.05710    0.74    0.46
```

```

## Year2000      0.06260      0.05471      1.14      0.25
## Year2001      0.01330      0.05873      0.23      0.82
## Year2002     -0.01909      0.05150     -0.37      0.71
## Year2003     -0.02400      0.05414     -0.44      0.66
## Year2004     -0.00115      0.05627     -0.02      0.98
## Year2005     -0.05245      0.05547     -0.95      0.34
## Year2006     -0.05419      0.05348     -1.01      0.31
## Year2007     -0.01902      0.05465     -0.35      0.73
## Year2008      0.00329      0.05369      0.06      0.95
## Year2009      0.00979      0.05377      0.18      0.86
## Year2010     -0.01904      0.05403     -0.35      0.72
## Year2011     -0.02476      0.05362     -0.46      0.64
## Year2012     -0.04650      0.05542     -0.84      0.40
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.35
## Multiple R-squared:  0.0494, Adjusted R-squared:  0.0397
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## observation 1810 is an outlier with |weight| = 0 ( < 4.6e-05);
## 170 weights are ~= 1. The remaining 2009 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0752 0.8630 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      4.59e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500      50      2      1      1000      200
## trace.lev      mts      compute.rd
##      0      1000      0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.110 1      1.054
## LastAuthorFemale 1.071 1      1.035
## Year      1.134 16      1.004

```

Residuals from first and last author



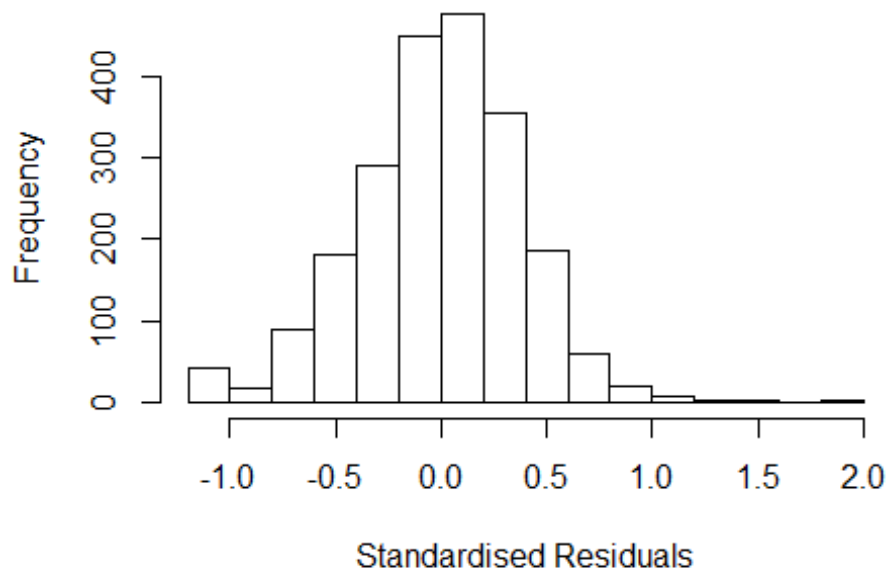
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.09965 -0.23424 0.00996 0.23918 1.96857
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.04702 0.04238 24.70 <2e-16 ***
## FirstAuthorFemale1 0.00159 0.01754 0.09 0.93
## LastAuthorFemale1 -0.02366 0.01869 -1.27 0.21
## Year1997 0.02558 0.05869 0.44 0.66
## Year1998 0.01232 0.05934 0.21 0.84
## Year1999 0.05263 0.05605 0.94 0.35
## Year2000 0.08644 0.05394 1.60 0.11
## Year2001 0.02156 0.05802 0.37 0.71
## Year2002 -0.00768 0.05012 -0.15 0.88
## Year2003 -0.00111 0.05280 -0.02 0.98
## Year2004 0.02647 0.05524 0.48 0.63
## Year2005 -0.03102 0.05393 -0.58 0.57
```

```

## Year2006      -0.02876    0.05235   -0.55    0.58
## Year2007      0.01665    0.05224    0.32    0.75
## Year2008      0.04415    0.05180    0.85    0.39
## Year2009      0.02280    0.05242    0.43    0.66
## Year2010      0.01741    0.05261    0.33    0.74
## Year2011      0.02037    0.05181    0.39    0.69
## Year2012     -0.00201    0.05353   -0.04    0.97
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.00586,    Adjusted R-squared:  -0.00242
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## observation 1810 is an outlier with |weight| = 0 ( < 4.6e-05);
## 178 weights are ~= 1. The remaining 2001 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0333 0.8650 0.9520 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      4.59e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.088 1          1.043
## Year              1.088 16          1.003

```

Residuals from first author

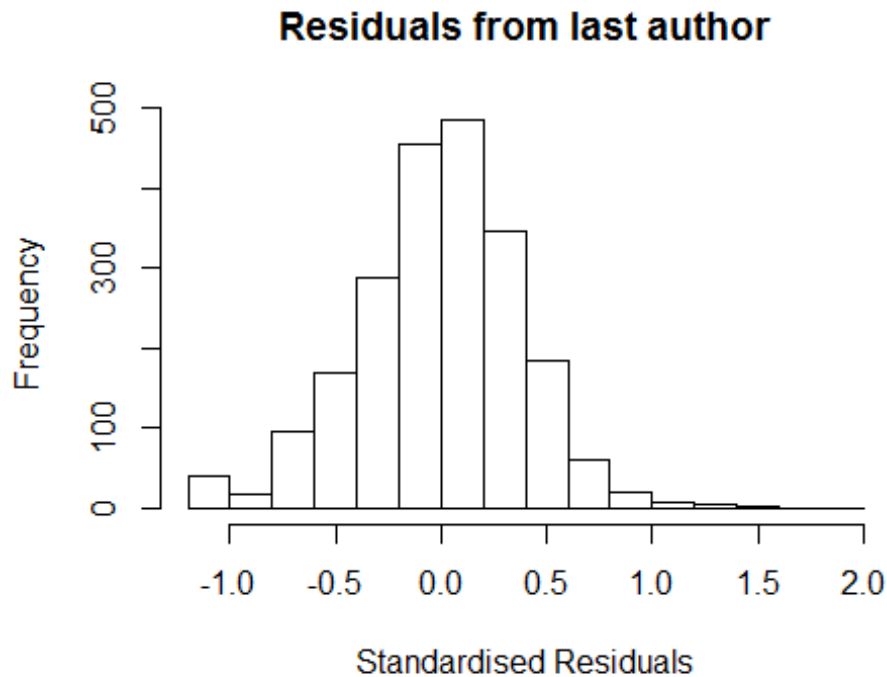


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0953 -0.2330 0.0112 0.2401 1.9730
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.04315 0.04247 24.56 <2e-16 ***
## FirstAuthorFemale1 -0.00237 0.01739 -0.14 0.89
## Year1997 0.02204 0.05865 0.38 0.71
## Year1998 0.01274 0.05952 0.21 0.83
## Year1999 0.05215 0.05634 0.93 0.35
## Year2000 0.08591 0.05410 1.59 0.11
## Year2001 0.02176 0.05827 0.37 0.71
## Year2002 -0.00659 0.05033 -0.13 0.90
## Year2003 -0.00208 0.05300 -0.04 0.97
## Year2004 0.02558 0.05533 0.46 0.64
## Year2005 -0.03063 0.05411 -0.57 0.57
## Year2006 -0.02794 0.05258 -0.53 0.60
```

```

## Year2007          0.01754    0.05252    0.33    0.74
## Year2008          0.04378    0.05206    0.84    0.40
## Year2009          0.02207    0.05266    0.42    0.68
## Year2010          0.01682    0.05287    0.32    0.75
## Year2011          0.01891    0.05201    0.36    0.72
## Year2012         -0.00175    0.05374   -0.03    0.97
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.359
## Multiple R-squared:  0.00518,    Adjusted R-squared:  -0.00264
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## observation 1810 is an outlier with |weight| = 0 ( < 4.6e-05);
## 170 weights are ~= 1. The remaining 2009 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0335 0.8650 0.9520 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      4.59e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## 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 0 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
## -1.1000 -0.2341  0.0101  0.2390  1.9679
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.04733    0.04229   24.77  <2e-16 ***
## LastAuthorFemale1 -0.02333    0.01851   -1.26    0.21
## Year1997         0.02559    0.05869    0.44    0.66
## Year1998         0.01232    0.05934    0.21    0.84
## Year1999         0.05266    0.05606    0.94    0.35
## Year2000         0.08627    0.05391    1.60    0.11
## Year2001         0.02162    0.05801    0.37    0.71
## Year2002        -0.00759    0.05010   -0.15    0.88
## Year2003        -0.00106    0.05280   -0.02    0.98
## Year2004         0.02666    0.05516    0.48    0.63
## Year2005        -0.03096    0.05393   -0.57    0.57
## Year2006        -0.02860    0.05229   -0.55    0.58
```

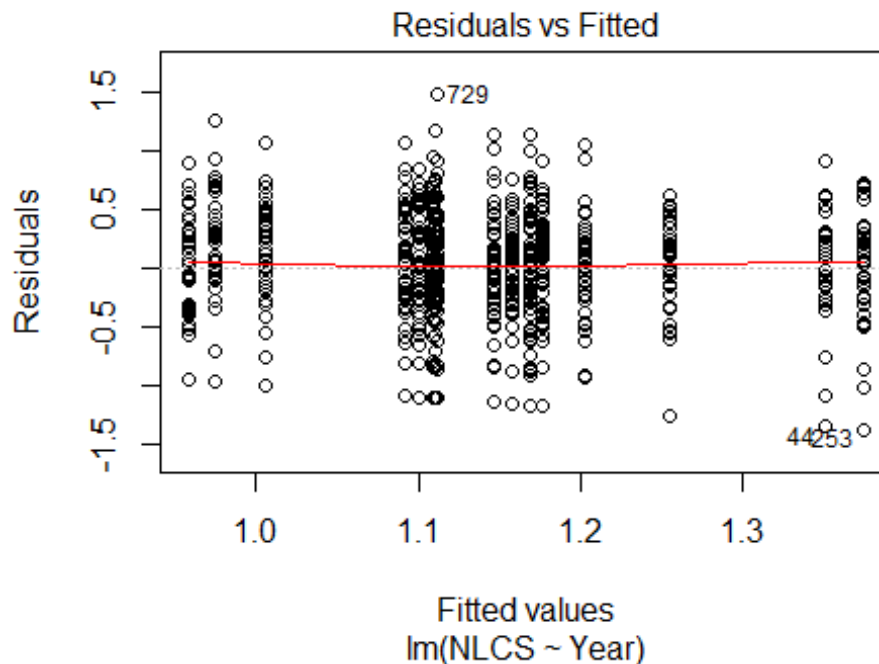
```

## Year2007      0.01697      0.05206      0.33      0.74
## Year2008      0.04446      0.05156      0.86      0.39
## Year2009      0.02295      0.05241      0.44      0.66
## Year2010      0.01773      0.05235      0.34      0.73
## Year2011      0.02065      0.05164      0.40      0.69
## Year2012     -0.00172      0.05323     -0.03      0.97
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.00586,    Adjusted R-squared:  -0.00195
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## observation 1810 is an outlier with |weight| = 0 ( < 4.6e-05);
## 176 weights are ~= 1. The remaining 2003 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0332 0.8660 0.9520 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      4.59e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 2180"
## [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
##   81   97   85   79  108  104   98   87   97   70   90  116  109  117  118
## 2011 2012
##  104  124
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   34   50   35   37   38   37   35   40   46   34   42   55   49   59   57

```



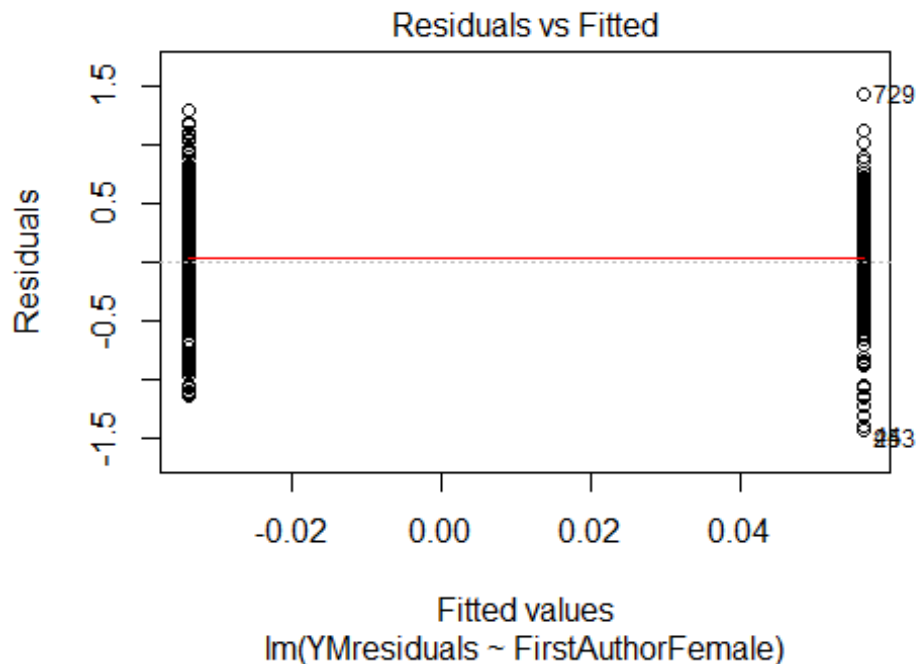
```
## 2011 2012
## 48 55
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 31 42 26 31 30 32 26 33 42 26 34 49 39 46 45
## 2011 2012
## 40 47
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



```
##
## 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: 4.48886574068587"
## [1] "Male first author team size 2018 geometric mean: 3.96547489072713"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

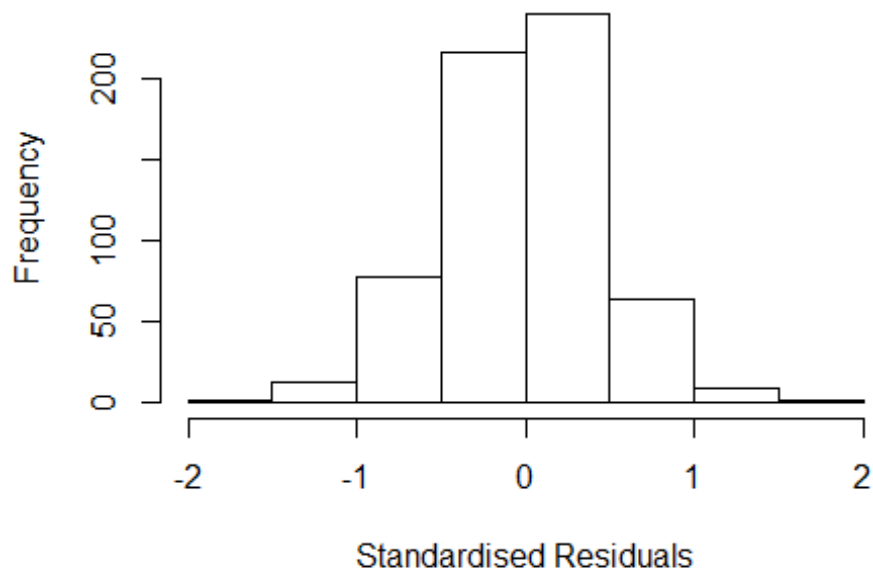
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 530, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.45894970383705"
## [1] "Male last author team size 2018 geometric mean: 4.08885290433978"

## 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 = 470, 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 1.289  1      1.135
## LastAuthorFemale  1.127  1      1.062
## UniqueAuthors    1.964  4      1.088
## Year              2.254 16      1.026
```

Residuals from first and last author and team size



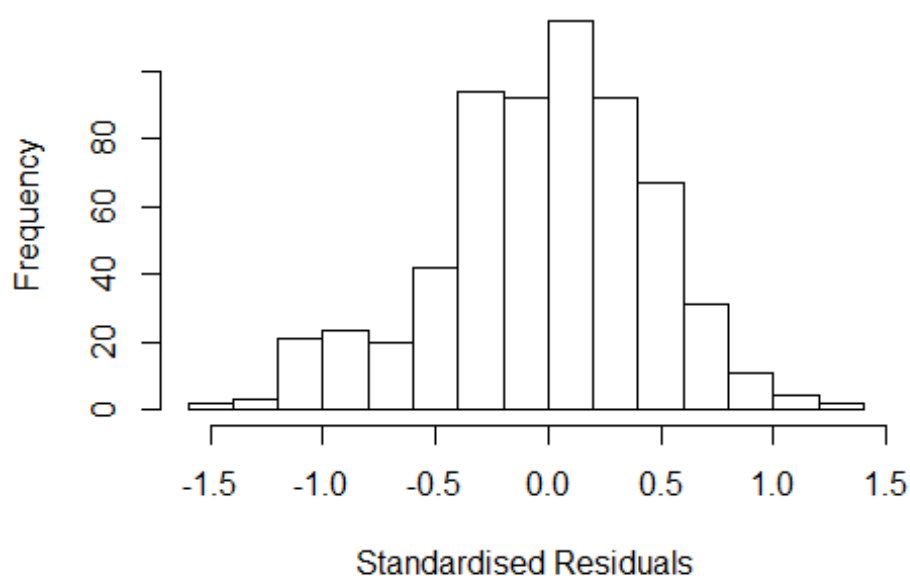
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.65657 -0.26311  0.00756  0.28706  1.69884
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0209    0.1104   9.25 < 2e-16 ***
## FirstAuthorFemale1  0.1098    0.0405   2.71  0.00684 **
## LastAuthorFemale1 -0.0432    0.0436  -0.99  0.32205
## UniqueAuthors2     0.3451    0.0872   3.96  8.4e-05 ***
## UniqueAuthors3     0.4116    0.0826   4.98  8.2e-07 ***
## UniqueAuthors4     0.3488    0.0873   4.00  7.2e-05 ***
## UniqueAuthors5     0.4650    0.0831   5.59  3.4e-08 ***
## Year1997          -0.1412    0.0992  -1.42  0.15502
## Year1998           0.0609    0.1220   0.50  0.61807
## Year1999          -0.2075    0.1099  -1.89  0.05944 .
```

```

## Year2000          -0.4888      0.1402    -3.49  0.00052 ***
## Year2001          -0.1956      0.1165    -1.68  0.09363 .
## Year2002          -0.2452      0.1258    -1.95  0.05173 .
## Year2003          -0.4198      0.1031    -4.07  5.3e-05 ***
## Year2004          -0.3645      0.1080    -3.38  0.00078 ***
## Year2005          -0.1386      0.1157    -1.20  0.23150
## Year2006          -0.2760      0.1219    -2.26  0.02387 *
## Year2007          -0.3134      0.1050    -2.98  0.00296 **
## Year2008          -0.2699      0.1045    -2.58  0.01005 *
## Year2009          -0.2373      0.0894    -2.65  0.00817 **
## Year2010          -0.3281      0.1087    -3.02  0.00266 **
## Year2011          -0.3258      0.1046    -3.11  0.00194 **
## Year2012          -0.2786      0.0983    -2.84  0.00474 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.434
## Multiple R-squared:  0.177, Adjusted R-squared:  0.147
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 60 weights are ~= 1. The remaining 559 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.091  0.868  0.952   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          1.62e-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 1.213 1          1.102
## LastAuthorFemale  1.104 1          1.051
## Year              1.300 16          1.008

```

Residuals from first and last author



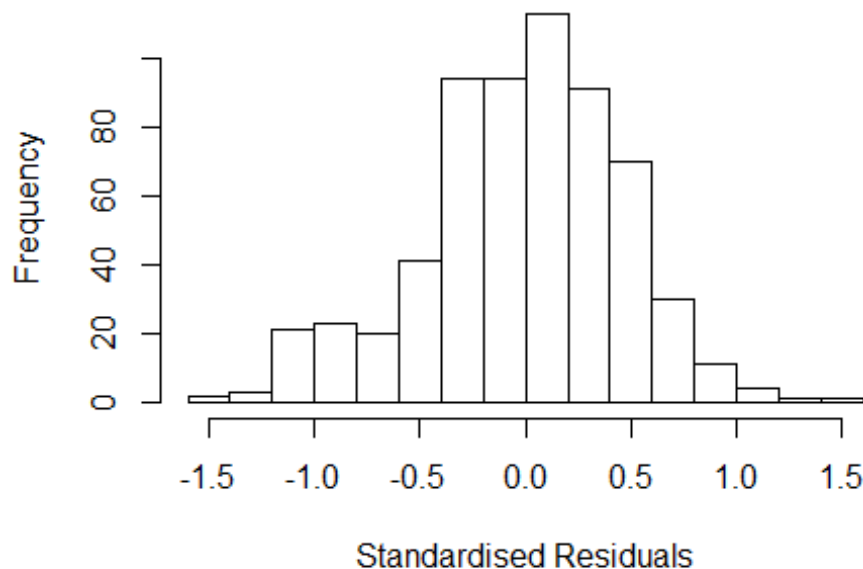
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5801 -0.2943 0.0158 0.2815 1.3937
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3954 0.0731 19.09 < 2e-16 ***
## FirstAuthorFemale1 0.1353 0.0402 3.36 0.00083 ***
## LastAuthorFemale1 -0.0376 0.0447 -0.84 0.40032
## Year1997 -0.2112 0.0939 -2.25 0.02482 *
## Year1998 0.0494 0.1171 0.42 0.67341
## Year1999 -0.3030 0.1105 -2.74 0.00627 **
## Year2000 -0.5581 0.1534 -3.64 0.00030 ***
## Year2001 -0.2719 0.1200 -2.27 0.02385 *
## Year2002 -0.2554 0.1202 -2.13 0.03392 *
## Year2003 -0.5472 0.1066 -5.13 3.9e-07 ***
## Year2004 -0.3801 0.1137 -3.34 0.00088 ***
## Year2005 -0.1789 0.1125 -1.59 0.11250
```

```

## Year2006          -0.3177      0.1206    -2.63  0.00865 **
## Year2007          -0.3353      0.1025    -3.27  0.00113 **
## Year2008          -0.3050      0.1015    -3.01  0.00276 **
## Year2009          -0.2579      0.0879    -2.93  0.00347 **
## Year2010          -0.3327      0.1063    -3.13  0.00184 **
## Year2011          -0.3453      0.1046    -3.30  0.00102 **
## Year2012          -0.2621      0.0958    -2.74  0.00640 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.0933, Adjusted R-squared:  0.0661
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 50 weights are ~= 1. The remaining 569 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.178  0.870  0.957  0.897  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      1.62e-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 1.198 1          1.095
## Year              1.198 16          1.006

```

Residuals from first author



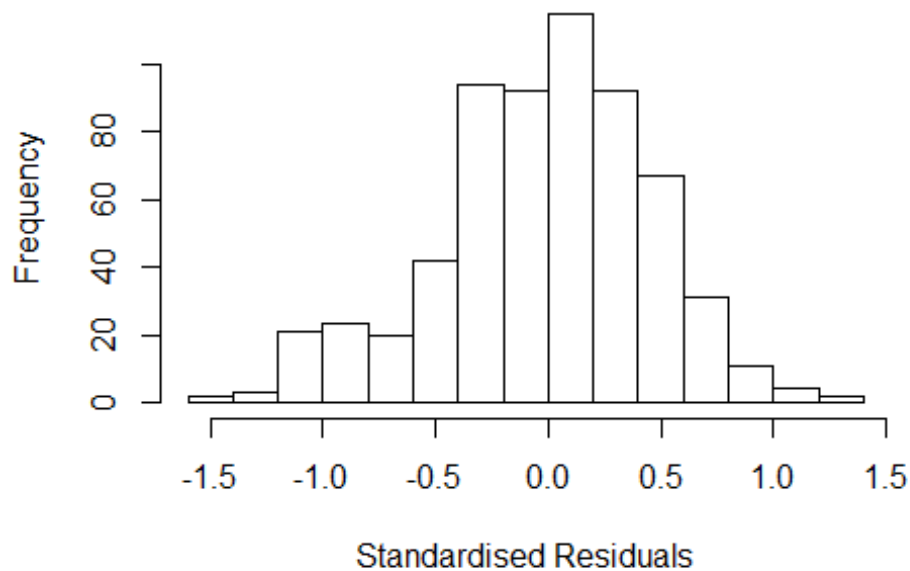
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5686 -0.2967 0.0218 0.2820 1.4015
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3873 0.0743 18.66 < 2e-16 ***
## FirstAuthorFemale1 0.1323 0.0400 3.31 0.00100 **
## Year1997 -0.2109 0.0945 -2.23 0.02593 *
## Year1998 0.0490 0.1188 0.41 0.68036
## Year1999 -0.3012 0.1110 -2.71 0.00684 **
## Year2000 -0.5577 0.1528 -3.65 0.00028 ***
## Year2001 -0.2741 0.1199 -2.29 0.02264 *
## Year2002 -0.2538 0.1205 -2.11 0.03562 *
## Year2003 -0.5480 0.1068 -5.13 3.8e-07 ***
## Year2004 -0.3763 0.1151 -3.27 0.00114 **
## Year2005 -0.1804 0.1131 -1.60 0.11117
## Year2006 -0.3153 0.1205 -2.62 0.00909 **
```

```

## Year2007          -0.3371      0.1033   -3.26  0.00116 **
## Year2008          -0.3107      0.1025   -3.03  0.00253 **
## Year2009          -0.2567      0.0890   -2.89  0.00404 **
## Year2010          -0.3314      0.1075   -3.08  0.00214 **
## Year2011          -0.3488      0.1057   -3.30  0.00102 **
## Year2012          -0.2654      0.0971   -2.73  0.00643 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.0926, Adjusted R-squared:  0.067
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 44 weights are ~= 1. The remaining 575 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.183  0.868  0.955  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      1.62e-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 1.103 1          1.050
## Year            1.103 16          1.003

```


Residuals from last author



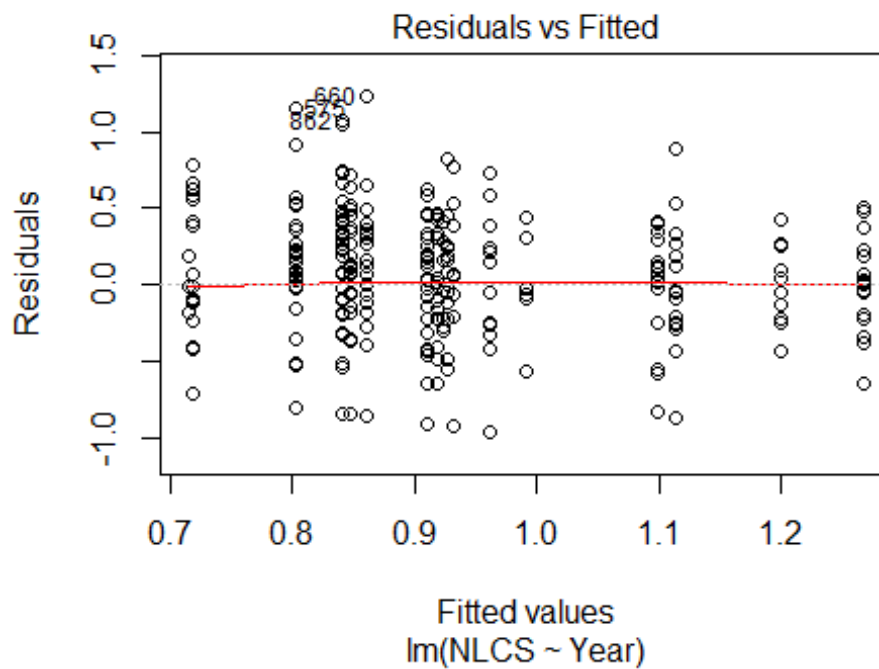
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4594 -0.2888 0.0183 0.2966 1.3878
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4341 0.0768 18.67 < 2e-16 ***
## LastAuthorFemale1 -0.0253 0.0450 -0.56 0.57397
## Year1997 -0.2369 0.0975 -2.43 0.01539 *
## Year1998 0.0253 0.1201 0.21 0.83322
## Year1999 -0.2977 0.1159 -2.57 0.01048 *
## Year2000 -0.5359 0.1610 -3.33 0.00093 ***
## Year2001 -0.2913 0.1247 -2.34 0.01978 *
## Year2002 -0.2299 0.1190 -1.93 0.05376 .
## Year2003 -0.5525 0.1102 -5.01 7.1e-07 ***
## Year2004 -0.3655 0.1166 -3.13 0.00181 **
## Year2005 -0.1565 0.1117 -1.40 0.16187
## Year2006 -0.3137 0.1278 -2.45 0.01441 *
```

```

## Year2007          -0.3179      0.1063    -2.99  0.00289 **
## Year2008          -0.2977      0.1039    -2.87  0.00430 **
## Year2009          -0.2418      0.0914    -2.65  0.00835 **
## Year2010          -0.3026      0.1068    -2.83  0.00475 **
## Year2011          -0.3107      0.1074    -2.89  0.00394 **
## Year2012          -0.2280      0.0995    -2.29  0.02231 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.0732, Adjusted R-squared:  0.047
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 56 weights are ~= 1. The remaining 563 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.255  0.867  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      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: 619"
## [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
##   27   38   25   45   34   53   43   39   41   19   25   48   41   49   29
## 2011 2012
##   41   61
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13    6    7   16    3   10   16   16   16   14   11   32   27   31   20
## 2011 2012

```

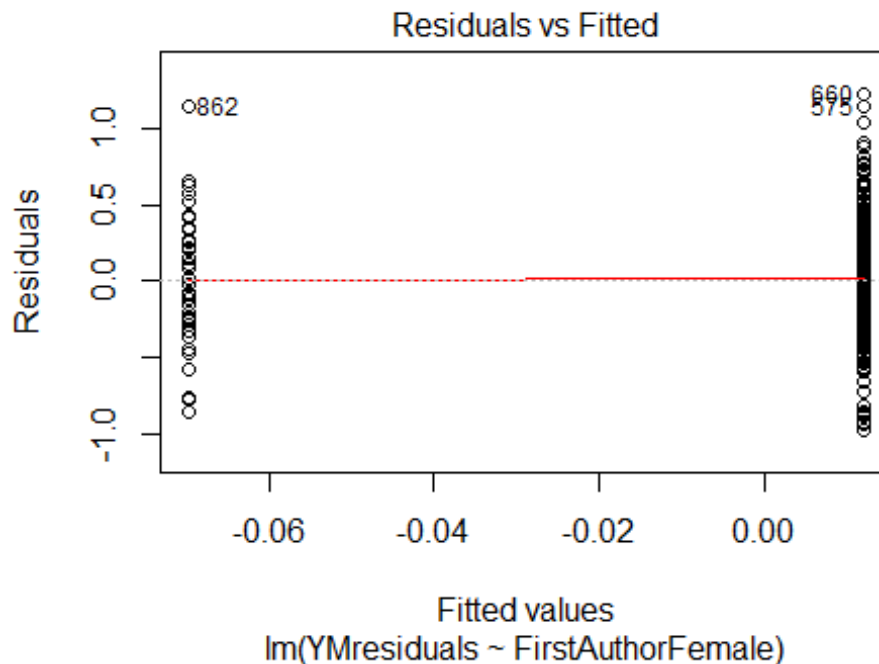
```
## 27 47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 6 7 14 3 10 14 14 14 8 9 29 21 29 18
## 2011 2012
## 24 44
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.079, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 3.87298334620742"
## [1] "Male first author team size 2018 geometric mean: 3.56328512953488"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

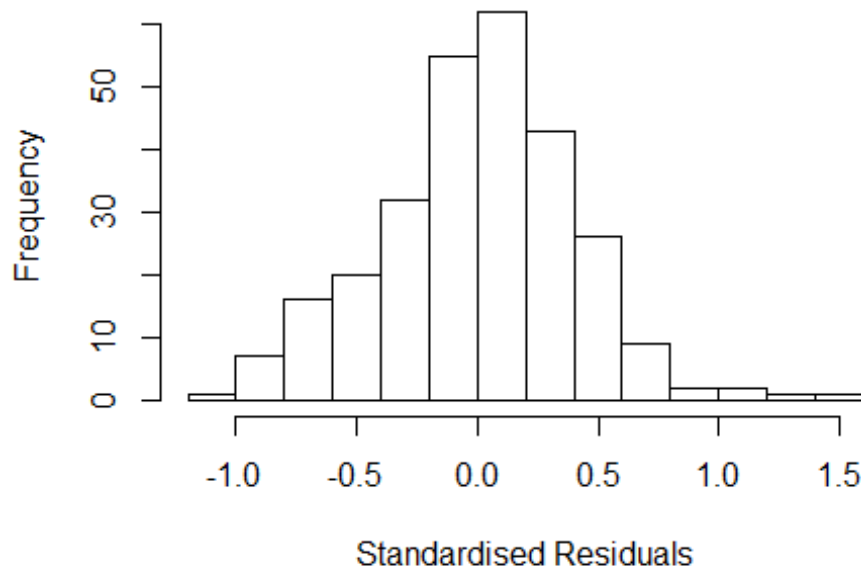
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.94923200383977"
## [1] "Male last author team size 2018 geometric mean: 3.21029709880649"

## 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 = 36, 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 1.276 1      1.130
## LastAuthorFemale  1.746 1      1.321
## UniqueAuthors    3.593 4      1.173
## Year              6.503 16     1.060
```

Residuals from first and last author and team size



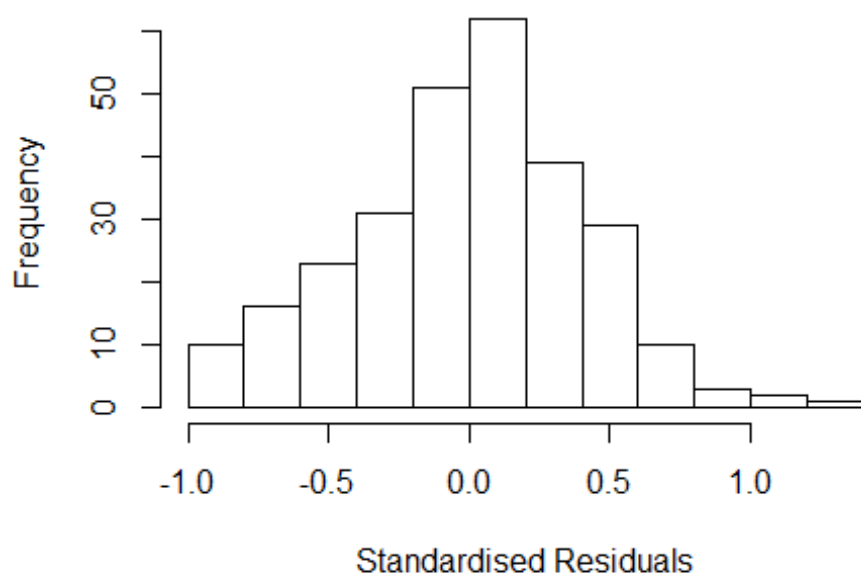
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0524 -0.2400 0.0176 0.2526 1.4785
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8448 0.1078 7.84 1.2e-13 ***
## FirstAuthorFemale1 -0.0622 0.0643 -0.97 0.33413
## LastAuthorFemale1 -0.0985 0.0657 -1.50 0.13499
## UniqueAuthors2 0.1964 0.0804 2.44 0.01524 *
## UniqueAuthors3 0.1829 0.0838 2.18 0.02994 *
## UniqueAuthors4 0.2658 0.0782 3.40 0.00078 ***
## UniqueAuthors5 0.3566 0.0875 4.08 6.1e-05 ***
## Year1997 0.0581 0.1731 0.34 0.73728
## Year1998 -0.0342 0.1366 -0.25 0.80260
## Year1999 -0.0211 0.1444 -0.15 0.88418
```

```

## Year2000          -0.1940      0.1510    -1.28  0.20006
## Year2001           0.2338      0.1632     1.43  0.15328
## Year2002           0.2910      0.1534     1.90  0.05893 .
## Year2003           0.1221      0.1406     0.87  0.38582
## Year2004           0.0750      0.1381     0.54  0.58744
## Year2005          -0.0858      0.1313    -0.65  0.51403
## Year2006          -0.1518      0.1979    -0.77  0.44357
## Year2007          -0.2340      0.1375    -1.70  0.08986 .
## Year2008          -0.2293      0.1375    -1.67  0.09656 .
## Year2009          -0.0614      0.1372    -0.45  0.65489
## Year2010          -0.3689      0.1463    -2.52  0.01226 *
## Year2011          -0.1349      0.1557    -0.87  0.38690
## Year2012          -0.1490      0.1426    -1.04  0.29712
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.206, Adjusted R-squared:  0.137
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 249 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.115  0.848  0.953  0.897  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      3.61e-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 1.229 1      1.109
## LastAuthorFemale  1.636 1      1.279
## Year              1.912 16      1.020

```

Residuals from first and last author



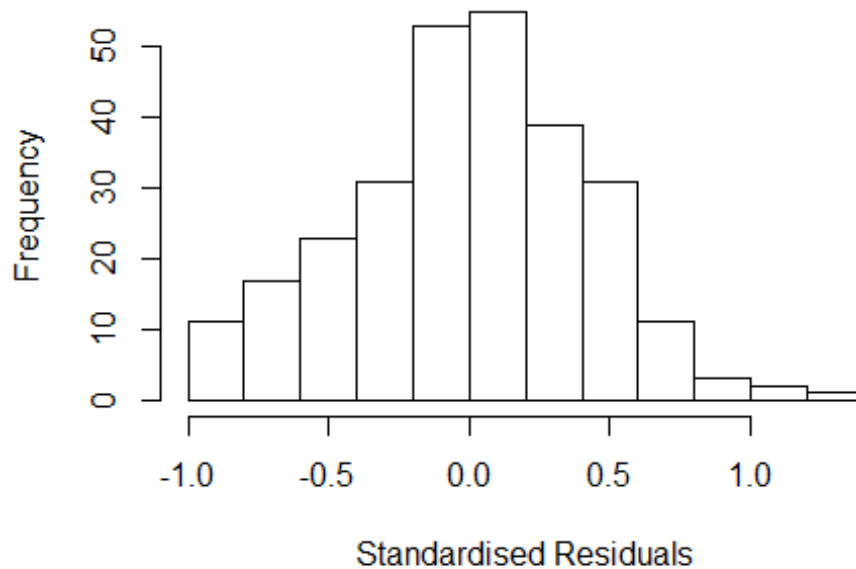
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9821 -0.2527 0.0361 0.2711 1.2957
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98211 0.12363 7.94 6.1e-14 ***
## FirstAuthorFemale1 -0.06652 0.06707 -0.99 0.322
## LastAuthorFemale1 -0.07168 0.06961 -1.03 0.304
## Year1997 0.02575 0.18974 0.14 0.892
## Year1998 -0.03257 0.16236 -0.20 0.841
## Year1999 -0.00157 0.15468 -0.01 0.992
## Year2000 -0.26798 0.15248 -1.76 0.080 .
## Year2001 0.21662 0.14951 1.45 0.149
## Year2002 0.30756 0.15815 1.94 0.053 .
## Year2003 0.17994 0.16334 1.10 0.272
## Year2004 0.11210 0.15770 0.71 0.478
## Year2005 -0.04205 0.15039 -0.28 0.780
```

```

## Year2006      -0.15117    0.20499   -0.74    0.462
## Year2007      -0.20724    0.16055   -1.29    0.198
## Year2008      -0.18385    0.15821   -1.16    0.246
## Year2009      -0.00351    0.14824   -0.02    0.981
## Year2010      -0.33240    0.17452   -1.90    0.058 .
## Year2011      -0.03823    0.15332   -0.25    0.803
## Year2012      -0.14017    0.15435   -0.91    0.365
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.134, Adjusted R-squared:  0.0731
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.297  0.862  0.956  0.904  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      3.61e-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 1.273 1      1.128
## Year              1.273 16      1.008

```


Residuals from first author



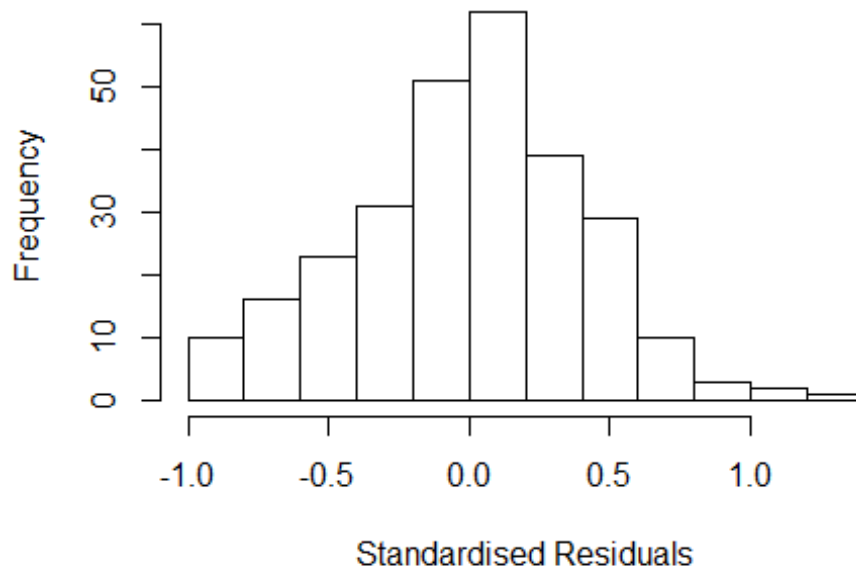
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9754 -0.2541 0.0139 0.2692 1.2996
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9754 0.1261 7.74 2.3e-13 ***
## FirstAuthorFemale1 -0.0908 0.0725 -1.25 0.212
## Year1997 0.0223 0.1853 0.12 0.904
## Year1998 -0.0295 0.1670 -0.18 0.860
## Year1999 -0.0116 0.1560 -0.07 0.941
## Year2000 -0.2613 0.1544 -1.69 0.092 .
## Year2001 0.2233 0.1514 1.47 0.141
## Year2002 0.3034 0.1580 1.92 0.056 .
## Year2003 0.1824 0.1670 1.09 0.276
## Year2004 0.1119 0.1604 0.70 0.486
## Year2005 -0.0451 0.1532 -0.29 0.769
## Year2006 -0.1503 0.2118 -0.71 0.479
```

```

## Year2007          -0.2085      0.1617    -1.29     0.198
## Year2008          -0.1810      0.1596    -1.13     0.258
## Year2009          -0.0187      0.1480    -0.13     0.899
## Year2010          -0.3331      0.1760    -1.89     0.059 .
## Year2011          -0.0593      0.1520    -0.39     0.697
## Year2012          -0.1470      0.1549    -0.95     0.343
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0704
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 252 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.325  0.878  0.953  0.907  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      3.61e-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 1.647 1      1.283
## Year      1.647 16      1.016

```

Residuals from last author



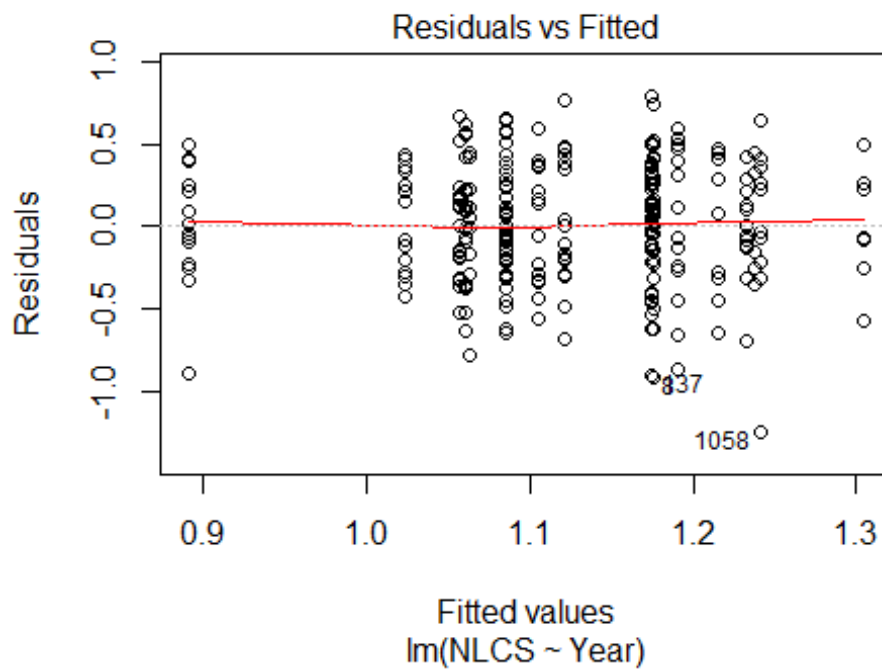
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9837 -0.2527 0.0235 0.2693 1.3024
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9837 0.1230 8.00 4.2e-14 ***
## LastAuthorFemale1 -0.0898 0.0726 -1.24 0.217
## Year1997 0.0265 0.1909 0.14 0.890
## Year1998 -0.0505 0.1593 -0.32 0.752
## Year1999 -0.0139 0.1532 -0.09 0.928
## Year2000 -0.2695 0.1519 -1.77 0.077 .
## Year2001 0.2151 0.1490 1.44 0.150
## Year2002 0.3010 0.1569 1.92 0.056 .
## Year2003 0.1744 0.1613 1.08 0.281
## Year2004 0.1054 0.1577 0.67 0.504
## Year2005 -0.0411 0.1497 -0.27 0.784
## Year2006 -0.1541 0.2066 -0.75 0.456
```

```

## Year2007          -0.2150      0.1587   -1.35    0.177
## Year2008          -0.1921      0.1581   -1.21    0.226
## Year2009          -0.0166      0.1468   -0.11    0.910
## Year2010          -0.3367      0.1734   -1.94    0.053 .
## Year2011          -0.0486      0.1520   -0.32    0.750
## Year2012          -0.1516      0.1529   -0.99    0.322
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.132, Adjusted R-squared:  0.075
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 253 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.291  0.867   0.955   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      3.61e-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: 277"
## [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
##   88   94   73   70   60   58   61   56   46   40   39   81   92   82   65
## 2011 2012
##   69   27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   23   38   19   17   21   14   20   20   16   14   16    9    9    5   11
## 2011 2012

```

```
## 15 7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 21 36 17 14 15 12 18 15 15 12 14 9 7 5 11
## 2011 2012
## 15 7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 11, df = 16, p-value = 0.8
```



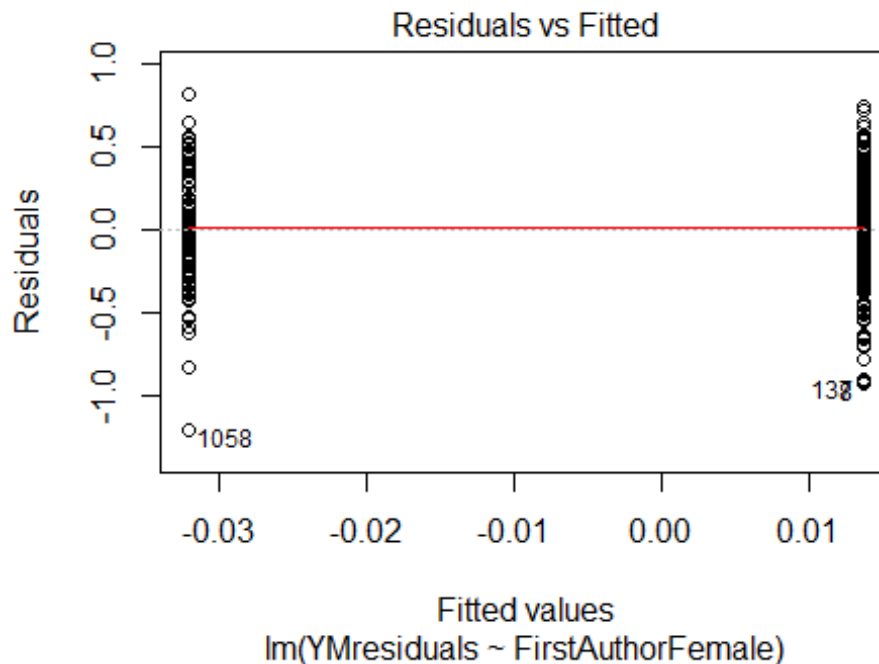
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.066, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 4.11215430794245"
## [1] "Male first author team size 2018 geometric mean: 4.76798920689919"

## 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 = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.37090777138021"
## [1] "Male last author team size 2018 geometric mean: 4.49499843757614"

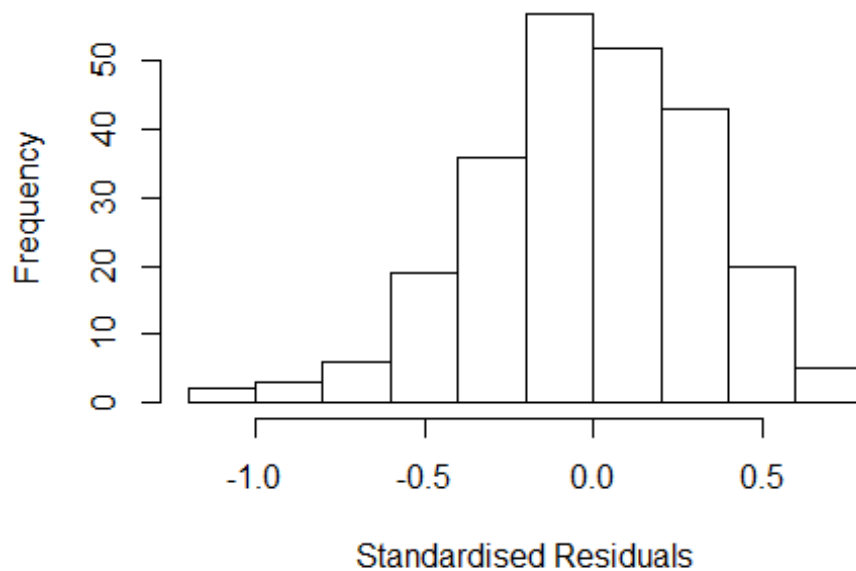
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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	1.527	1	1.236
LastAuthorFemale	1.626	1	1.275
UniqueAuthors	3.732	4	1.179
Year	6.987	16	1.063

Residuals from first and last author and team size



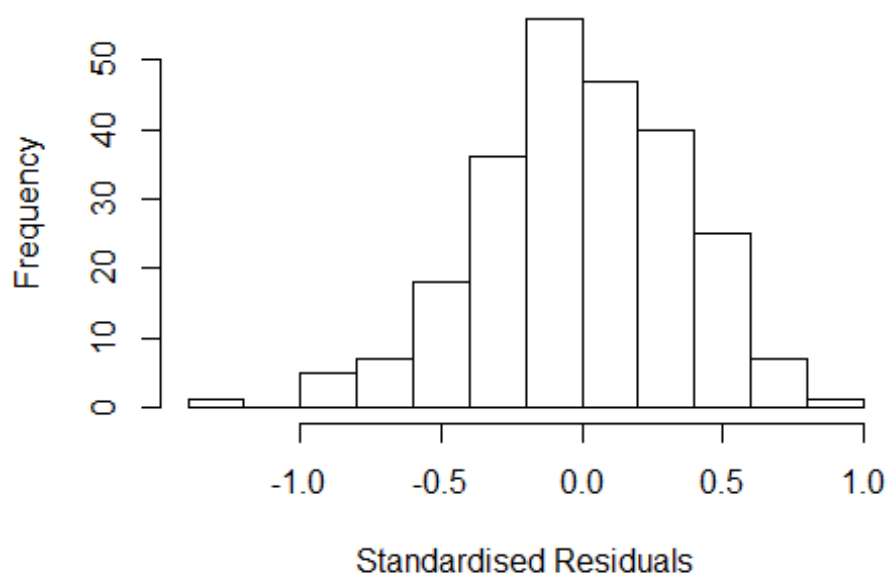
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.092 -0.209 -0.011 0.224 0.775
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06804 0.08933 11.96 < 2e-16 ***
## FirstAuthorFemale1 -0.10071 0.05317 -1.89 0.05952 .
## LastAuthorFemale1 0.00194 0.05423 0.04 0.97151
## UniqueAuthors2 0.04024 0.07896 0.51 0.61082
## UniqueAuthors3 0.14266 0.07607 1.88 0.06205 .
## UniqueAuthors4 0.19522 0.09322 2.09 0.03738 *
## UniqueAuthors5 0.28154 0.07811 3.60 0.00039 ***
## Year1997 0.02470 0.10146 0.24 0.80790
## Year1998 -0.10841 0.13033 -0.83 0.40641
## Year1999 0.07323 0.14249 0.51 0.60780
```

```

## Year2000      -0.03732    0.12291   -0.30  0.76168
## Year2001      0.06698    0.11198    0.60  0.55040
## Year2002     -0.13642    0.11405   -1.20  0.23290
## Year2003     -0.15750    0.11164   -1.41  0.15974
## Year2004     -0.14789    0.13607   -1.09  0.27828
## Year2005     -0.13991    0.13669   -1.02  0.30716
## Year2006     -0.35997    0.11017   -3.27  0.00126 **
## Year2007      0.05155    0.16819    0.31  0.75951
## Year2008     -0.23953    0.15254   -1.57  0.11778
## Year2009     -0.03208    0.15562   -0.21  0.83689
## Year2010      0.08402    0.16058    0.52  0.60134
## Year2011     -0.01742    0.16580   -0.11  0.91643
## Year2012      0.14484    0.12929    1.12  0.26381
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.341
## Multiple R-squared:  0.163, Adjusted R-squared:  0.0797
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 225 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.285  0.879  0.959  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      4.12e-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 1.487 1      1.219
## LastAuthorFemale  1.318 1      1.148
## Year              1.669 16      1.016

```


Residuals from first and last author



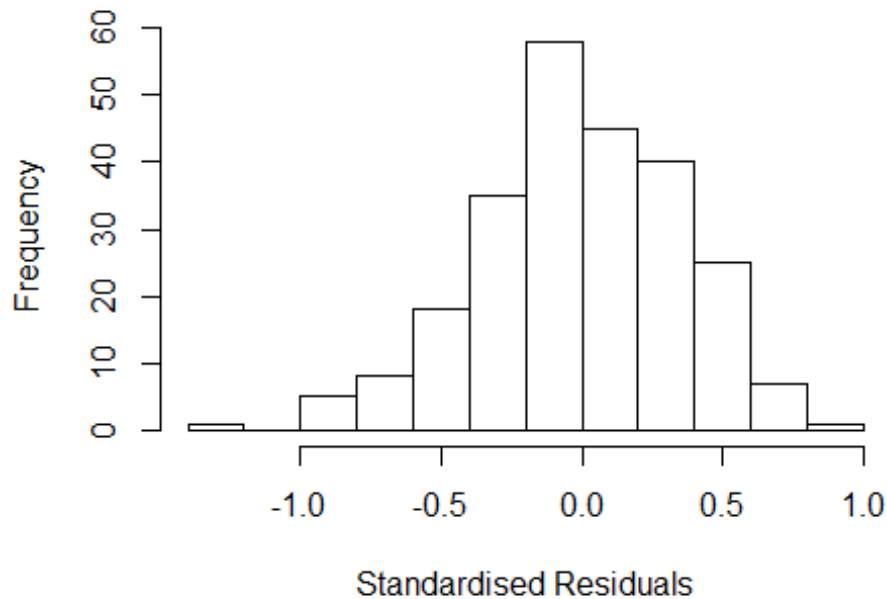
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2761 -0.2427 -0.0136 0.2445 0.8375
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.16759 0.08864 13.17 <2e-16 ***
## FirstAuthorFemale1 -0.09582 0.05419 -1.77 0.078 .
## LastAuthorFemale1 -0.01167 0.05537 -0.21 0.833
## Year1997 0.05268 0.10394 0.51 0.613
## Year1998 -0.10161 0.13404 -0.76 0.449
## Year1999 0.08322 0.14816 0.56 0.575
## Year2000 -0.00229 0.12210 -0.02 0.985
## Year2001 0.06800 0.11474 0.59 0.554
## Year2002 -0.09963 0.11407 -0.87 0.383
## Year2003 -0.11309 0.11624 -0.97 0.332
## Year2004 -0.05939 0.11943 -0.50 0.619
## Year2005 -0.10411 0.14094 -0.74 0.461
```

```

## Year2006      -0.29079    0.11272   -2.58    0.011 *
## Year2007      0.11282    0.18477    0.61    0.542
## Year2008     -0.17666    0.15860   -1.11    0.267
## Year2009      0.08436    0.17941    0.47    0.639
## Year2010      0.20435    0.13531    1.51    0.132
## Year2011      0.08095    0.16126    0.50    0.616
## Year2012      0.18258    0.15268    1.20    0.233
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.109, Adjusted R-squared:  0.037
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 224 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.183  0.879  0.951  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      4.12e-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 1.355 1      1.164
## Year              1.355 16      1.010

```

Residuals from first author



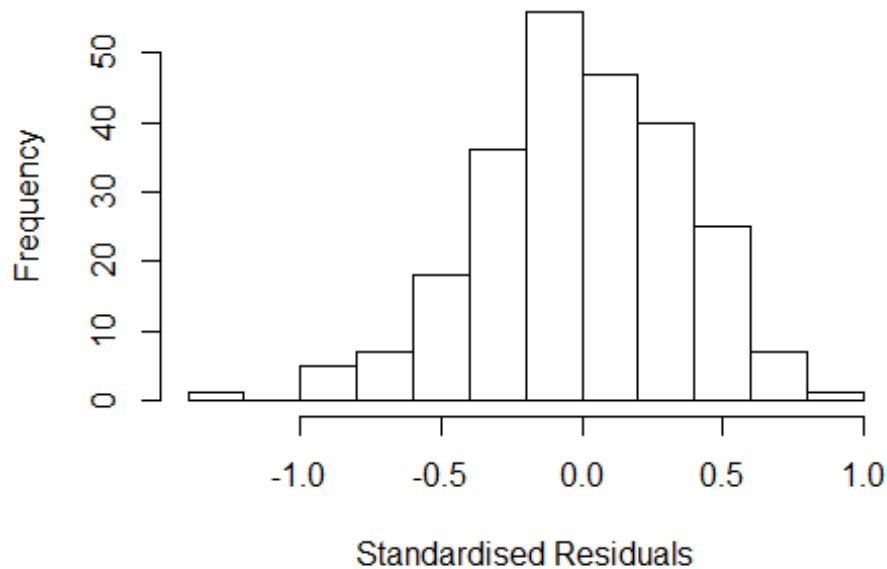
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2701 -0.2420 -0.0163 0.2464 0.8424
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.16650 0.08882 13.13 <2e-16 ***
## FirstAuthorFemale1 -0.09698 0.05306 -1.83 0.0689 .
## Year1997 0.05008 0.10354 0.48 0.6291
## Year1998 -0.10406 0.13337 -0.78 0.4361
## Year1999 0.08402 0.14806 0.57 0.5710
## Year2000 -0.00188 0.12212 -0.02 0.9877
## Year2001 0.06838 0.11492 0.60 0.5524
## Year2002 -0.10045 0.11410 -0.88 0.3796
## Year2003 -0.11412 0.11595 -0.98 0.3261
## Year2004 -0.05985 0.12016 -0.50 0.6189
## Year2005 -0.10615 0.14087 -0.75 0.4519
## Year2006 -0.29258 0.11247 -2.60 0.0099 **
```

```

## Year2007          0.11005    0.18476    0.60    0.5520
## Year2008         -0.17724    0.15937   -1.11    0.2673
## Year2009          0.08383    0.17804    0.47    0.6382
## Year2010          0.20055    0.13424    1.49    0.1366
## Year2011          0.07994    0.16149    0.50    0.6211
## Year2012          0.18193    0.15308    1.19    0.2359
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.109, Adjusted R-squared:  0.0417
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 224 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.182  0.879   0.951   0.912   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      4.12e-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          1.094
## Year            1.198 16          1.006

```

Residuals from last author



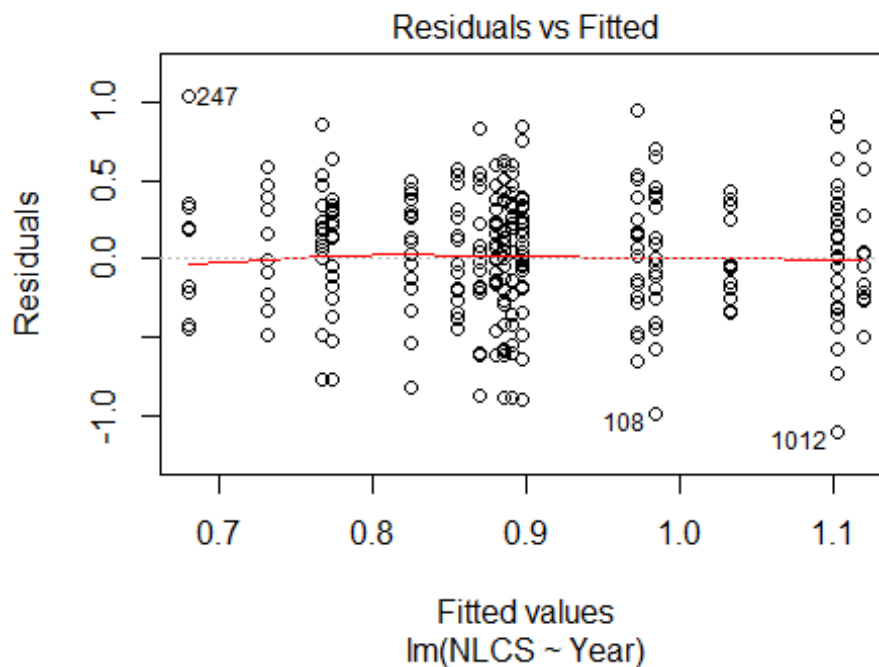
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.34928 -0.23785 0.00325 0.24866 0.75588
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1431 0.0858 13.32 <2e-16 ***
## LastAuthorFemale1 -0.0214 0.0552 -0.39 0.70
## Year1997 0.0630 0.1028 0.61 0.54
## Year1998 -0.0869 0.1322 -0.66 0.51
## Year1999 0.0678 0.1453 0.47 0.64
## Year2000 -0.0135 0.1227 -0.11 0.91
## Year2001 0.0750 0.1124 0.67 0.51
## Year2002 -0.1098 0.1122 -0.98 0.33
## Year2003 -0.1123 0.1152 -0.97 0.33
## Year2004 -0.0985 0.1186 -0.83 0.41
## Year2005 -0.1069 0.1385 -0.77 0.44
## Year2006 -0.2887 0.1119 -2.58 0.01 *
```

```

## Year2007          0.0938      0.1871      0.50      0.62
## Year2008         -0.1514      0.1563     -0.97      0.33
## Year2009          0.0888      0.1877      0.47      0.64
## Year2010          0.2062      0.1394      1.48      0.14
## Year2011          0.0758      0.1605      0.47      0.64
## Year2012          0.1714      0.1584      1.08      0.28
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.367
## Multiple R-squared:  0.0971, Adjusted R-squared:  0.0289
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 222 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.148  0.876  0.953  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      4.12e-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: 243"
## [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
##   33  42  35  33  41  34  30  27  37  30  35  77  99  85  69
## 2011 2012
##   74  50
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12  23  21  12  14  14  13  17  25  18  22  20  29  17  20
## 2011 2012

```

```
## 22 25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 23 21 12 14 13 12 15 19 18 18 20 26 16 20
## 2011 2012
## 22 21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 8.3, df = 16, p-value = 0.9
```



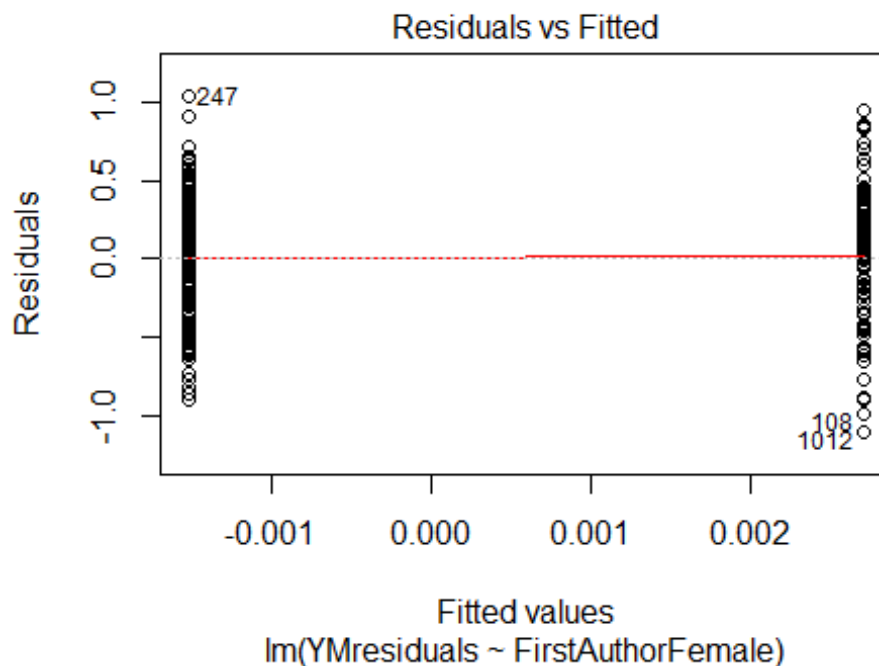
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.6, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 4.0404164254556"
## [1] "Male first author team size 2018 geometric mean: 4.24246714432811"

## 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 = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.18464316406821"
## [1] "Male last author team size 2018 geometric mean: 4.1097242634108"

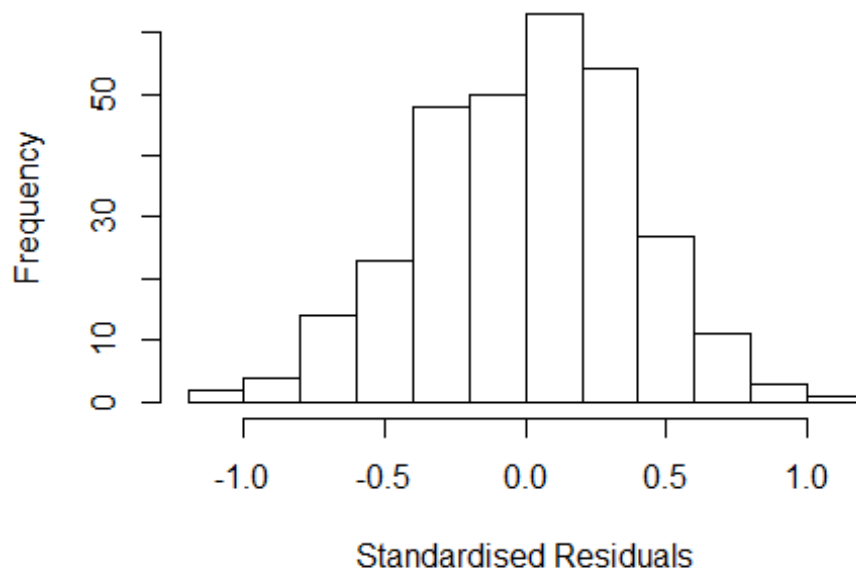
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 260, 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	1.514	1	1.231
LastAuthorFemale	1.412	1	1.188
UniqueAuthors	2.587	4	1.126
Year	3.856	16	1.043

Residuals from first and last author and team size



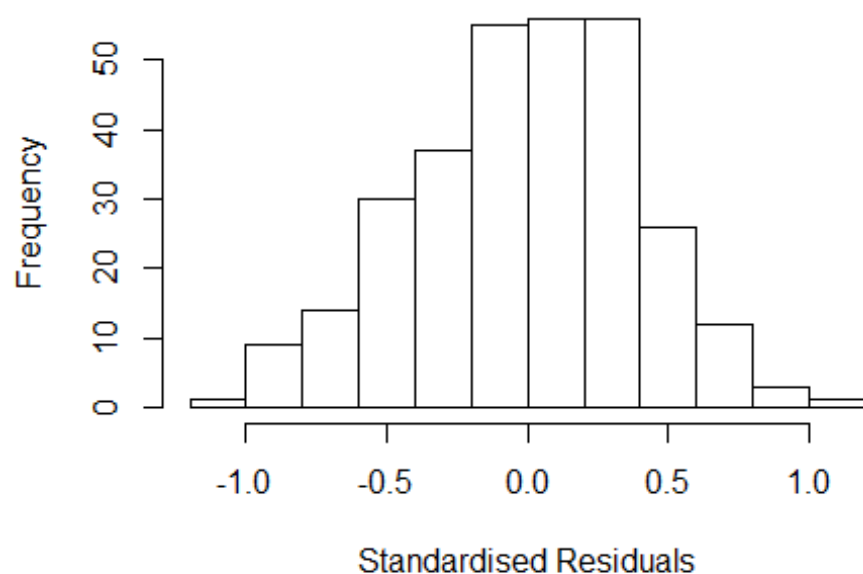
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1595 -0.2876 0.0214 0.2702 1.1434
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9859 0.0804 12.27 <2e-16 ***
## FirstAuthorFemale1 0.0507 0.0585 0.87 0.3874
## LastAuthorFemale1 -0.1140 0.0574 -1.98 0.0482 *
## UniqueAuthors2 0.1519 0.0695 2.19 0.0296 *
## UniqueAuthors3 0.2019 0.0773 2.61 0.0095 **
## UniqueAuthors4 0.2602 0.0934 2.79 0.0057 **
## UniqueAuthors5 0.2569 0.0932 2.76 0.0062 **
## Year1997 -0.1773 0.0968 -1.83 0.0682 .
## Year1998 -0.0915 0.1196 -0.76 0.4451
## Year1999 -0.0317 0.1418 -0.22 0.8235
```

```

## Year2000          -0.3369      0.1105    -3.05    0.0025 **
## Year2001          -0.4114      0.1371    -3.00    0.0029 **
## Year2002          -0.1937      0.1368    -1.42    0.1579
## Year2003          -0.2585      0.1144    -2.26    0.0246 *
## Year2004          -0.3952      0.1282    -3.08    0.0023 **
## Year2005          -0.2477      0.1263    -1.96    0.0510 .
## Year2006          -0.2787      0.1145    -2.43    0.0155 *
## Year2007          -0.3501      0.1253    -2.79    0.0056 **
## Year2008          -0.2602      0.1114    -2.34    0.0202 *
## Year2009          -0.2954      0.1475    -2.00    0.0461 *
## Year2010          -0.1224      0.1192    -1.03    0.3053
## Year2011          -0.2616      0.1487    -1.76    0.0796 .
## Year2012          -0.0200      0.1352    -0.15    0.8826
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.144, Adjusted R-squared:  0.0759
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 272 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.370  0.879  0.947  0.909  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          3.33e-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 1.454 1          1.206
## LastAuthorFemale  1.471 1          1.213
## Year              1.618 16          1.015

```

Residuals from first and last author



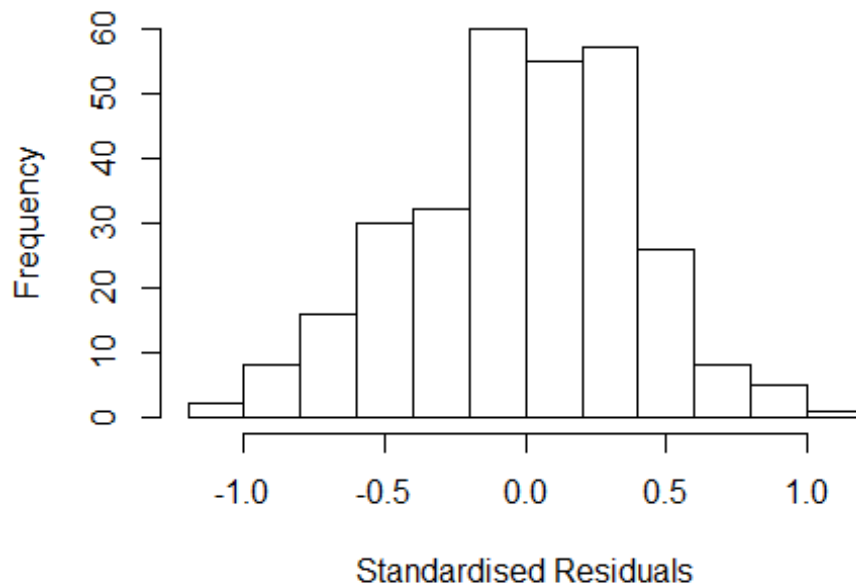
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0791 -0.2780 0.0241 0.2928 1.0370
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0726 0.0871 12.31 <2e-16 ***
## FirstAuthorFemale1 0.0511 0.0591 0.86 0.388
## LastAuthorFemale1 -0.1200 0.0590 -2.03 0.043 *
## Year1997 -0.1542 0.1122 -1.38 0.170
## Year1998 -0.0540 0.1254 -0.43 0.667
## Year1999 0.0440 0.1336 0.33 0.742
## Year2000 -0.3269 0.1350 -2.42 0.016 *
## Year2001 -0.3915 0.1534 -2.55 0.011 *
## Year2002 -0.1463 0.1636 -0.89 0.372
## Year2003 -0.2103 0.1266 -1.66 0.098 .
## Year2004 -0.3255 0.1312 -2.48 0.014 *
## Year2005 -0.1797 0.1359 -1.32 0.187
```

```

## Year2006          -0.1819      0.1219   -1.49    0.137
## Year2007          -0.2620      0.1351   -1.94    0.053 .
## Year2008          -0.1496      0.1180   -1.27    0.206
## Year2009          -0.2140      0.1458   -1.47    0.143
## Year2010          -0.0584      0.1348   -0.43    0.665
## Year2011          -0.1447      0.1382   -1.05    0.296
## Year2012           0.0754      0.1424    0.53    0.597
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.0983, Adjusted R-squared:  0.0405
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 272 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.441  0.865   0.945   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      3.33e-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 1.299 1      1.140
## Year              1.299 16      1.008

```

Residuals from first author



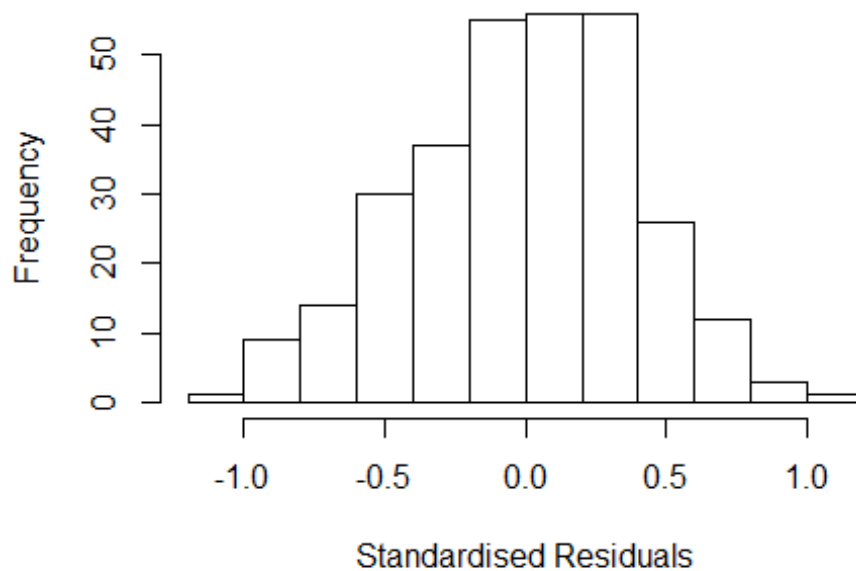
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1533 -0.2670 0.0086 0.2729 1.0678
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0677 0.0913 11.70 <2e-16 ***
## FirstAuthorFemale1 0.0111 0.0561 0.20 0.8427
## Year1997 -0.1846 0.1185 -1.56 0.1205
## Year1998 -0.0749 0.1295 -0.58 0.5632
## Year1999 0.0269 0.1360 0.20 0.8436
## Year2000 -0.3400 0.1380 -2.46 0.0143 *
## Year2001 -0.4175 0.1520 -2.75 0.0064 **
## Year2002 -0.1830 0.1648 -1.11 0.2679
## Year2003 -0.2292 0.1258 -1.82 0.0696 .
## Year2004 -0.3535 0.1358 -2.60 0.0097 **
## Year2005 -0.1949 0.1365 -1.43 0.1546
## Year2006 -0.1957 0.1236 -1.58 0.1143
```

```

## Year2007          -0.2895      0.1395   -2.08   0.0389 *
## Year2008          -0.1821      0.1224   -1.49   0.1380
## Year2009          -0.2182      0.1486   -1.47   0.1432
## Year2010          -0.1117      0.1308   -0.85   0.3939
## Year2011          -0.1661      0.1433   -1.16   0.2473
## Year2012           0.0745      0.1489    0.50   0.6172
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.0845, Adjusted R-squared:  0.0293
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 270 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.382  0.874   0.948   0.904   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      3.33e-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 1.253 1          1.119
## Year            1.253 16          1.007

```

Residuals from last author



```
## [1] "List of 0 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
## -1.0697 -0.2808 0.0188 0.2847 1.0300
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0808 0.0886 12.20 <2e-16 ***
## LastAuthorFemale1 -0.1030 0.0543 -1.90 0.0589 .
## Year1997 -0.1445 0.1133 -1.28 0.2033
## Year1998 -0.0519 0.1265 -0.41 0.6817
## Year1999 0.0449 0.1348 0.33 0.7392
## Year2000 -0.3261 0.1362 -2.39 0.0173 *
## Year2001 -0.3927 0.1505 -2.61 0.0095 **
## Year2002 -0.1499 0.1652 -0.91 0.3651
## Year2003 -0.2107 0.1251 -1.68 0.0933 .
## Year2004 -0.3232 0.1358 -2.38 0.0180 *
## Year2005 -0.1834 0.1366 -1.34 0.1803
## Year2006 -0.1770 0.1228 -1.44 0.1503
```

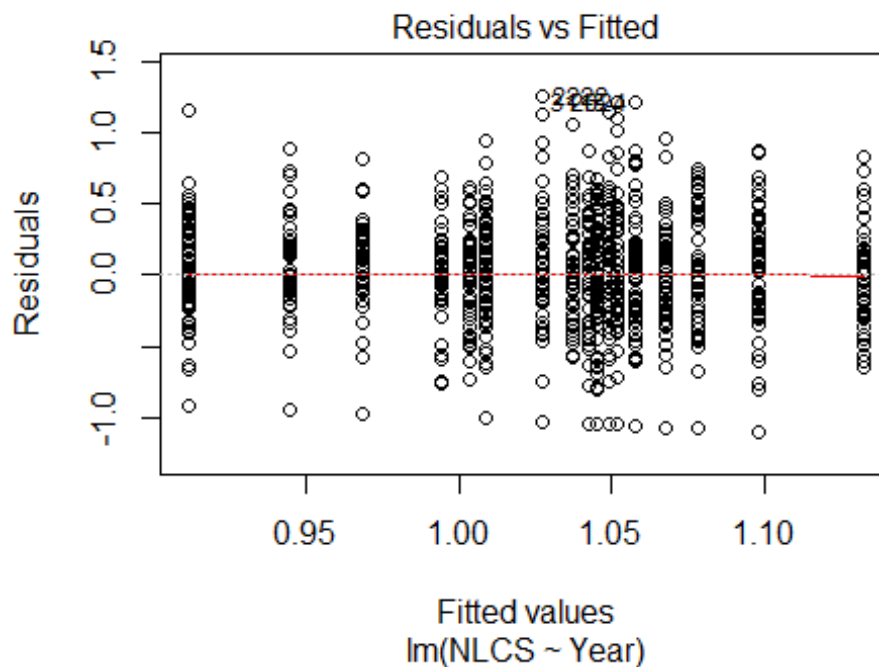
```

## Year2007          -0.2571      0.1364    -1.89    0.0604 .
## Year2008          -0.1492      0.1201    -1.24    0.2151
## Year2009          -0.2019      0.1454    -1.39    0.1660
## Year2010          -0.0593      0.1360    -0.44    0.6628
## Year2011          -0.1270      0.1366    -0.93    0.3533
## Year2012           0.0919      0.1406     0.65    0.5140
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.395
## Multiple R-squared:  0.0958, Adjusted R-squared:  0.0413
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 278 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.444  0.868  0.944  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      3.33e-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: 300"
## [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
##  179  204  208  194  173  158  168  160  157  153  114  115  115  132  116
## 2011 2012
##  122  129
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   51   56   59   55   48   44   68   63   53   70   46   47   46   57   52
## 2011 2012

```



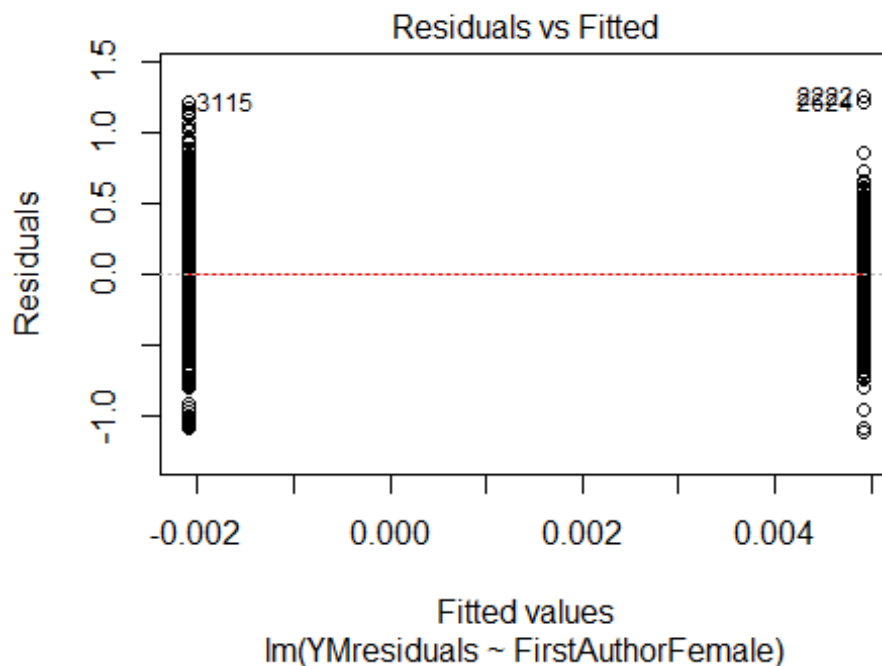
```
## 59 61
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 47 53 55 51 46 36 59 53 51 59 41 41 42 56 44
## 2011 2012
## 56 55
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 36, df = 16, p-value = 0.003
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.7, df = 1, p-value = 0.01
## [1] "Female first author team size 2018 geometric mean: 3.73178396993289"
## [1] "Male first author team size 2018 geometric mean: 3.76849956661992"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

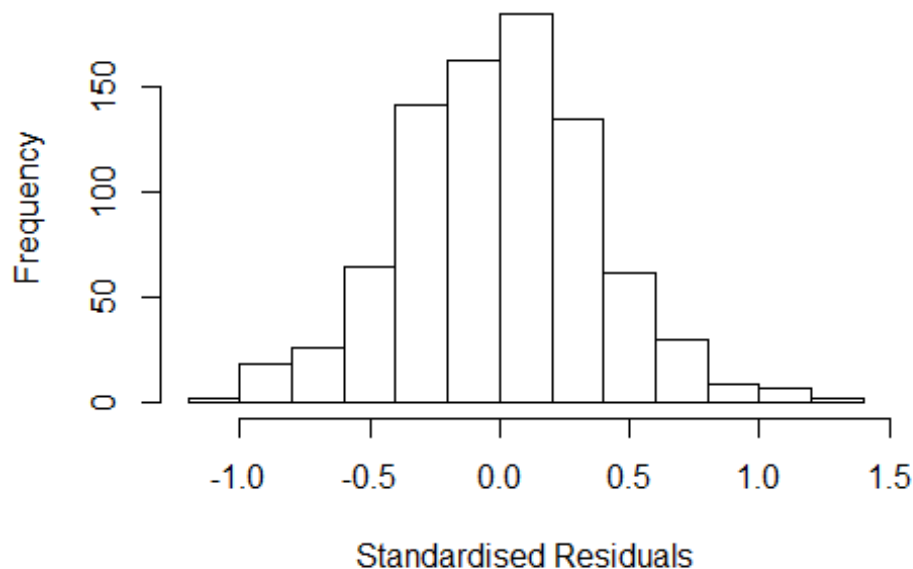
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 480, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46115652836729"
## [1] "Male last author team size 2018 geometric mean: 3.90429496505178"

## 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 = 0.3
## 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 1.177  1      1.085
## LastAuthorFemale  1.125  1      1.061
## UniqueAuthors    1.385  4      1.042
## Year              1.701 16      1.017
```

Residuals from first and last author and team size



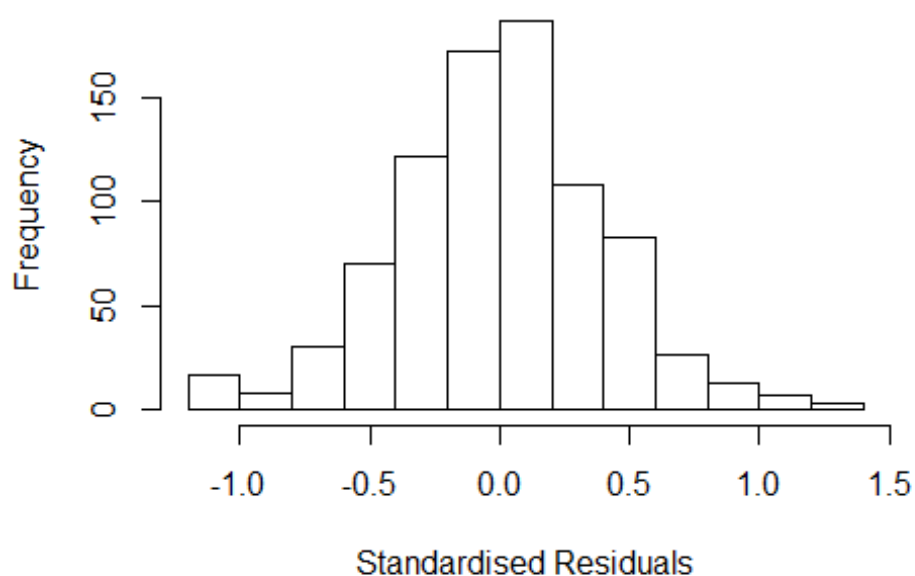
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0331 -0.2481  0.0106  0.2364  1.3277
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.84309    0.06414   13.14 < 2e-16 ***
## FirstAuthorFemale1 -0.00895    0.02918   -0.31  0.759
## LastAuthorFemale1 -0.01154    0.02945   -0.39  0.695
## UniqueAuthors2     0.19004    0.04364    4.35 1.5e-05 ***
## UniqueAuthors3     0.20482    0.04280    4.79 2.0e-06 ***
## UniqueAuthors4     0.33637    0.04586    7.33 5.4e-13 ***
## UniqueAuthors5     0.42308    0.04902    8.63 < 2e-16 ***
## Year1997           0.03101    0.07604    0.41  0.684
## Year1998          -0.00862    0.08357   -0.10  0.918
## Year1999          -0.02553    0.07865   -0.32  0.746
```

```

## Year2000      0.05763    0.07345    0.78    0.433
## Year2001     -0.06929    0.08981   -0.77    0.441
## Year2002      0.02292    0.07471    0.31    0.759
## Year2003     -0.08379    0.07379   -1.14    0.257
## Year2004     -0.05650    0.06936   -0.81    0.416
## Year2005     -0.16569    0.07228   -2.29    0.022 *
## Year2006     -0.11137    0.07760   -1.44    0.152
## Year2007     -0.08580    0.09622   -0.89    0.373
## Year2008     -0.03470    0.08712   -0.40    0.691
## Year2009     -0.07761    0.07961   -0.97    0.330
## Year2010     -0.01771    0.09412   -0.19    0.851
## Year2011      0.03218    0.07884    0.41    0.683
## Year2012     -0.10103    0.08757   -1.15    0.249
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.136, Adjusted R-squared:  0.113
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 66 weights are ~= 1. The remaining 779 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.145  0.877  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      1.18e-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 1.146 1      1.071
## LastAuthorFemale  1.104 1      1.051
## Year              1.255 16      1.007

```

Residuals from first and last author



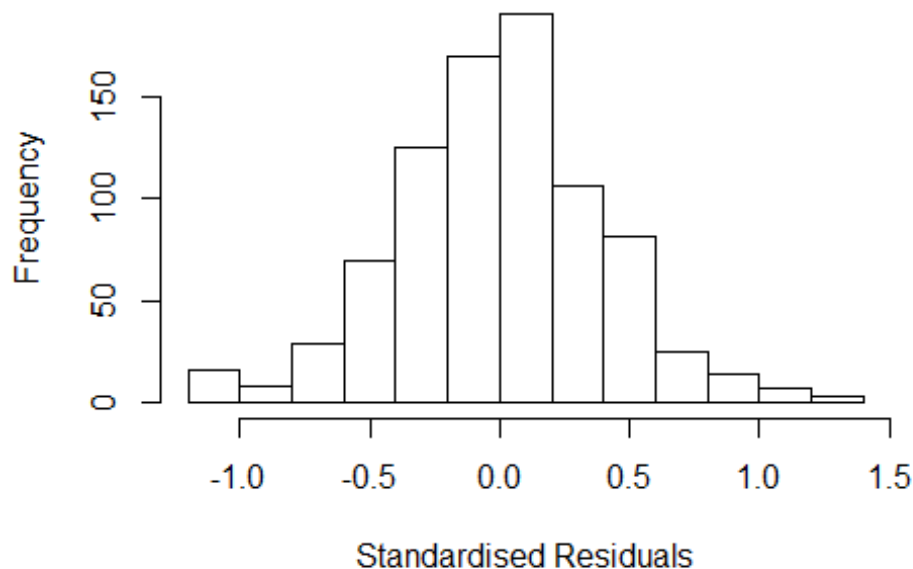
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.07375 -0.23748 0.00696 0.23463 1.27160
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02022 0.06228 16.38 <2e-16 ***
## FirstAuthorFemale1 0.01027 0.02988 0.34 0.731
## LastAuthorFemale1 0.01791 0.03066 0.58 0.559
## Year1997 0.03208 0.08016 0.40 0.689
## Year1998 0.04326 0.08755 0.49 0.621
## Year1999 -0.02799 0.08235 -0.34 0.734
## Year2000 0.10591 0.07917 1.34 0.181
## Year2001 -0.05466 0.09043 -0.60 0.546
## Year2002 0.03315 0.08026 0.41 0.680
## Year2003 -0.06324 0.07757 -0.82 0.415
## Year2004 -0.00261 0.07413 -0.04 0.972
## Year2005 -0.12696 0.07631 -1.66 0.097 .
```

```

## Year2006      -0.08129    0.08125   -1.00    0.317
## Year2007      -0.02008    0.10696   -0.19    0.851
## Year2008      -0.01634    0.08940   -0.18    0.855
## Year2009       0.01114    0.08494    0.13    0.896
## Year2010       0.04978    0.09789    0.51    0.611
## Year2011       0.08698    0.08345    1.04    0.298
## Year2012       0.00178    0.09373    0.02    0.985
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.376
## Multiple R-squared:  0.0232, Adjusted R-squared:  0.0019
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 49 weights are ~= 1. The remaining 796 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.230  0.863   0.956   0.900   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      1.18e-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 1.148 1         1.071
## Year              1.148 16         1.004

```

Residuals from first author



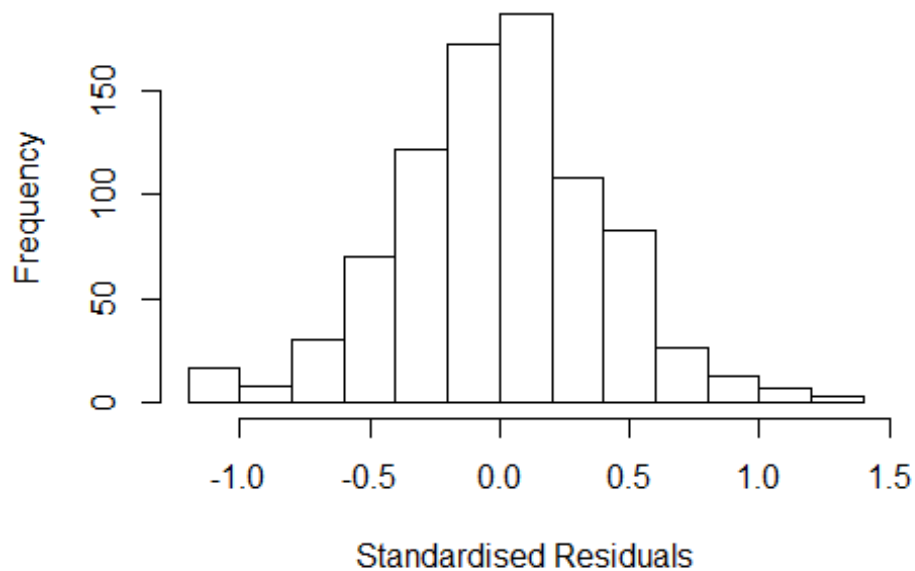
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.08078 -0.23961 0.00514 0.23174 1.26391
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02407 0.06124 16.72 <2e-16 ***
## FirstAuthorFemale1 0.01262 0.02991 0.42 0.673
## Year1997 0.03292 0.07983 0.41 0.680
## Year1998 0.04410 0.08709 0.51 0.613
## Year1999 -0.02816 0.08183 -0.34 0.731
## Year2000 0.10210 0.07806 1.31 0.191
## Year2001 -0.05443 0.08987 -0.61 0.545
## Year2002 0.03358 0.08014 0.42 0.675
## Year2003 -0.06298 0.07758 -0.81 0.417
## Year2004 -0.00137 0.07386 -0.02 0.985
## Year2005 -0.12800 0.07598 -1.68 0.092 .
## Year2006 -0.08082 0.08072 -1.00 0.317
```

```

## Year2007      -0.01860    0.10641   -0.17    0.861
## Year2008      -0.01589    0.08892   -0.18    0.858
## Year2009       0.01117    0.08445    0.13    0.895
## Year2010       0.05067    0.09741    0.52    0.603
## Year2011       0.08778    0.08288    1.06    0.290
## Year2012       0.00250    0.09333    0.03    0.979
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.376
## Multiple R-squared:  0.0228, Adjusted R-squared:  0.0027
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 47 weights are ~= 1. The remaining 798 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.235  0.863  0.958  0.900  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      1.18e-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 1.109 1      1.053
## Year              1.109 16      1.003

```


Residuals from last author



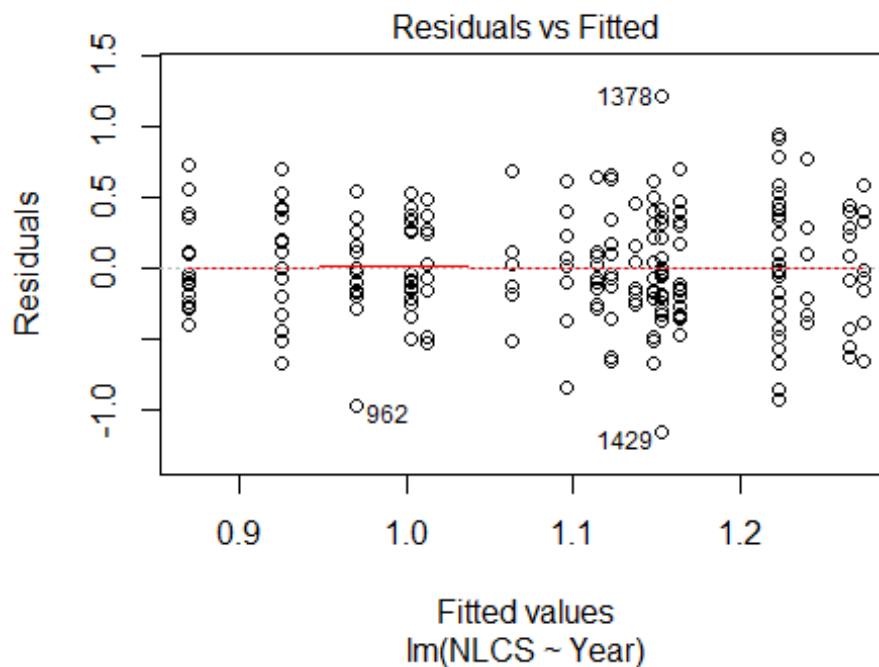
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.07384 -0.23558 0.00774 0.23849 1.27817
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.021415 0.062381 16.37 <2e-16 ***
## LastAuthorFemale1 0.019408 0.030755 0.63 0.53
## Year1997 0.034271 0.079930 0.43 0.67
## Year1998 0.044237 0.087376 0.51 0.61
## Year1999 -0.026099 0.082077 -0.32 0.75
## Year2000 0.107179 0.079122 1.35 0.18
## Year2001 -0.052274 0.090006 -0.58 0.56
## Year2002 0.034167 0.080228 0.43 0.67
## Year2003 -0.062652 0.077622 -0.81 0.42
## Year2004 -0.000817 0.073755 -0.01 0.99
## Year2005 -0.125315 0.076061 -1.65 0.10 .
## Year2006 -0.079582 0.080912 -0.98 0.33
```

```

## Year2007          -0.017585    0.105863    -0.17     0.87
## Year2008          -0.014619    0.089069    -0.16     0.87
## Year2009           0.011898    0.084954     0.14     0.89
## Year2010           0.052422    0.096979     0.54     0.59
## Year2011           0.088848    0.082836     1.07     0.28
## Year2012           0.004820    0.092601     0.05     0.96
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.377
## Multiple R-squared:  0.023, Adjusted R-squared:  0.00295
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 51 weights are ~= 1. The remaining 794 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.227  0.862  0.958  0.900  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      1.18e-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: 845"
## [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
##   80   84   93  103   94   68   73   66   72   78   55   60   66   47   52
## 2011 2012
##   60   75
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10    8    8    6    9   11   16    7    8   15   14   17   18   11   15
## 2011 2012

```

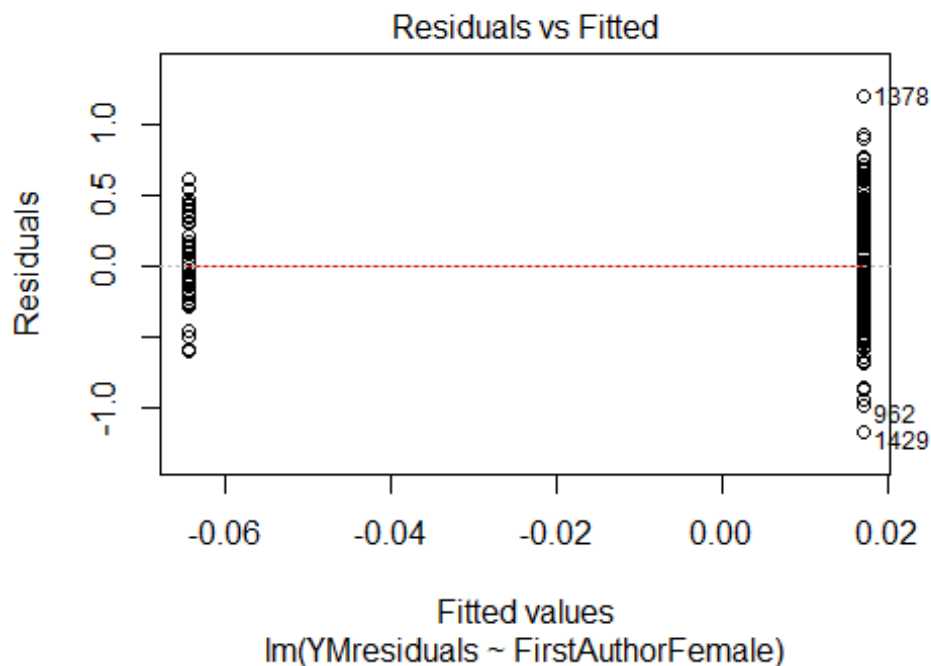
```
## 19 33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 6 7 4 8 11 12 5 8 13 12 15 16 11 12
## 2011 2012
## 19 31
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 12, df = 16, p-value = 0.7
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.2, df = 1, p-value = 0.01
## [1] "Female first author team size 2018 geometric mean: 4.29109355007577"
## [1] "Male first author team size 2018 geometric mean: 4.26755221621256"
## 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 = 84, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.57085181225251"
## [1] "Male last author team size 2018 geometric mean: 4.14911219969377"

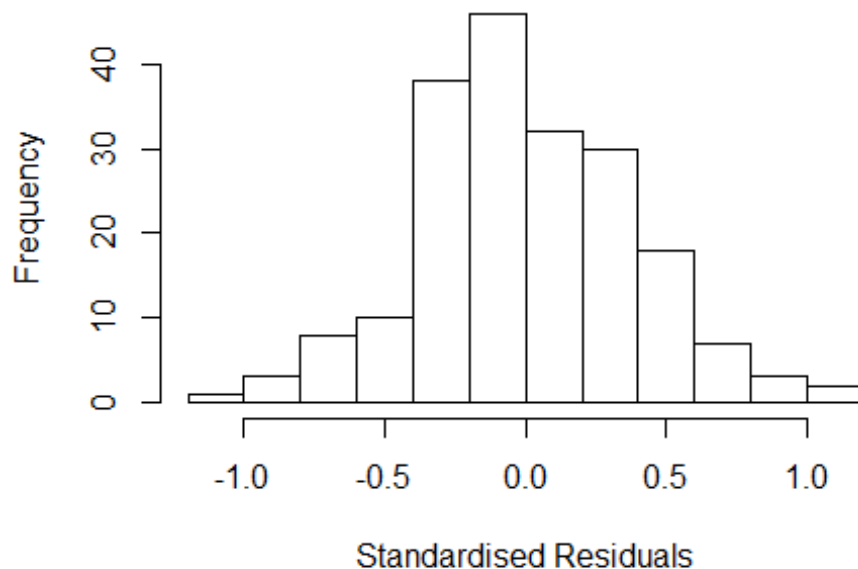
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 82, p-value = 0.7
## 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	2.151	1	1.467
LastAuthorFemale	1.630	1	1.277
UniqueAuthors	3.555	4	1.172
Year	7.479	16	1.065

Residuals from first and last author and team size



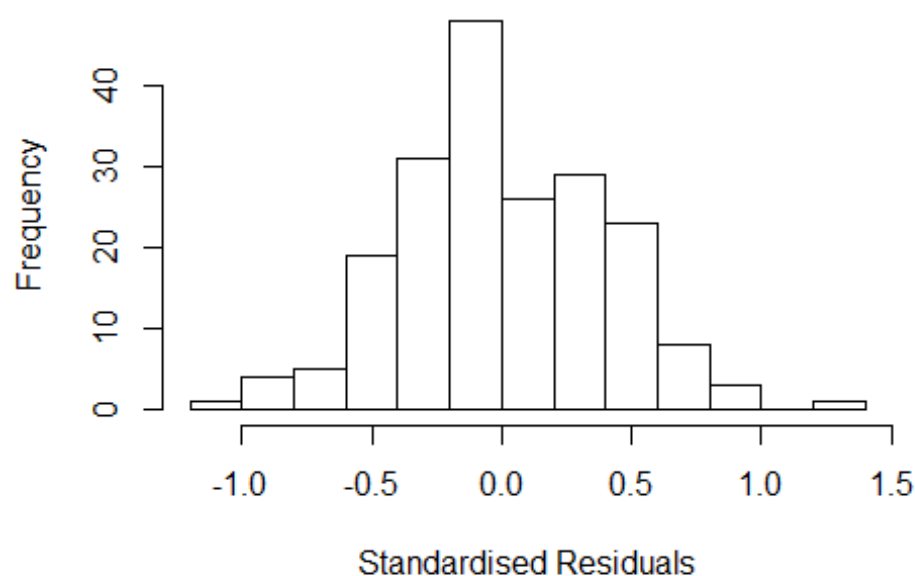
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1345 -0.2636 -0.0452 0.2951 1.0485
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.072748 0.164824 6.51 7.7e-10 ***
## FirstAuthorFemale1 -0.138930 0.080282 -1.73 0.085 .
## LastAuthorFemale1 -0.013698 0.089042 -0.15 0.878
## UniqueAuthors2 0.184217 0.116120 1.59 0.114
## UniqueAuthors3 0.099676 0.126931 0.79 0.433
## UniqueAuthors4 0.234068 0.132527 1.77 0.079 .
## UniqueAuthors5 0.279667 0.131010 2.13 0.034 *
## Year1997 -0.000441 0.176307 0.00 0.998
## Year1998 0.143649 0.235312 0.61 0.542
## Year1999 0.008175 0.202676 0.04 0.968
```

```

## Year2000      -0.167674    0.168232    -1.00    0.320
## Year2001      -0.068544    0.199311    -0.34    0.731
## Year2002      -0.256476    0.192005    -1.34    0.183
## Year2003      -0.033284    0.222010    -0.15    0.881
## Year2004      -0.076359    0.206712    -0.37    0.712
## Year2005      -0.292893    0.169730    -1.73    0.086 .
## Year2006      -0.094961    0.199631    -0.48    0.635
## Year2007      -0.209062    0.159496    -1.31    0.192
## Year2008      -0.326152    0.159651    -2.04    0.043 *
## Year2009      -0.194265    0.166687    -1.17    0.245
## Year2010      -0.058847    0.186517    -0.32    0.753
## Year2011      -0.037938    0.160931    -0.24    0.814
## Year2012      0.023800    0.168226    0.14    0.888
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.125, Adjusted R-squared:  0.0155
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 182 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.409  0.898  0.952  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      5.05e-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 1.711 1      1.308
## LastAuthorFemale  1.599 1      1.264
## Year              2.661 16      1.031

```

Residuals from first and last author



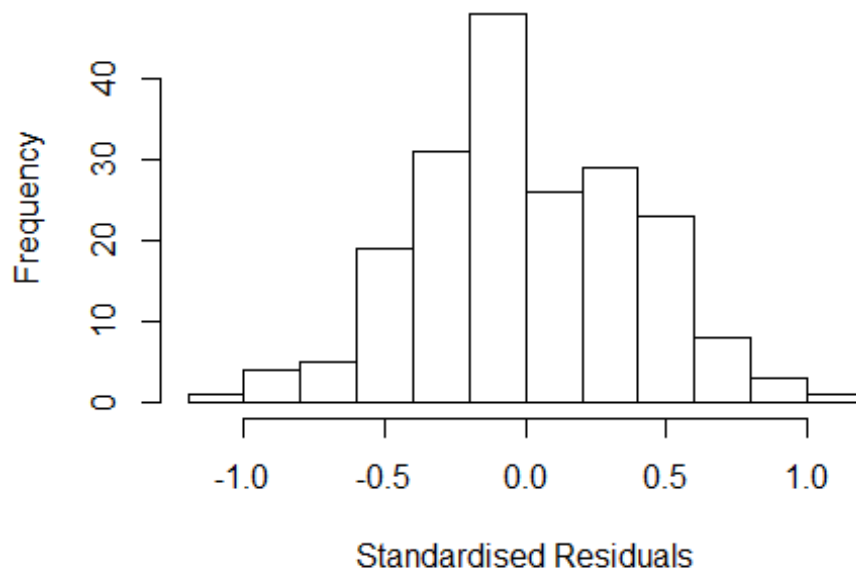
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1623 -0.2493 -0.0405  0.3223  1.2007
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21602    0.17150   7.09   3e-11 ***
## FirstAuthorFemale1 -0.07282    0.07322  -0.99    0.32
## LastAuthorFemale1  0.00773    0.09066   0.09    0.93
## Year1997        -0.03378    0.20274  -0.17    0.87
## Year1998         0.08889    0.24686   0.36    0.72
## Year1999        -0.05516    0.25363  -0.22    0.83
## Year2000        -0.12795    0.19358  -0.66    0.51
## Year2001        -0.08295    0.21954  -0.38    0.71
## Year2002        -0.27584    0.21337  -1.29    0.20
## Year2003        -0.01024    0.25276  -0.04    0.97
## Year2004        -0.07299    0.24186  -0.30    0.76
## Year2005        -0.25863    0.19501  -1.33    0.19
```

```

## Year2006      -0.07301    0.21503   -0.34    0.73
## Year2007      -0.18099    0.19167   -0.94    0.35
## Year2008      -0.31711    0.19456   -1.63    0.10
## Year2009      -0.17539    0.21296   -0.82    0.41
## Year2010      -0.05746    0.21687   -0.26    0.79
## Year2011      -0.05367    0.19321   -0.28    0.78
## Year2012       0.04356    0.20102    0.22    0.83
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.0843, Adjusted R-squared:  -0.00783
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 183 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.399  0.889  0.955  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      5.05e-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 1.732 1      1.316
## Year              1.732 16      1.017

```


Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.1640 -0.2501 -0.0386 0.3216 1.1990
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2163 0.1735 7.01 4.6e-11 ***
## FirstAuthorFemale1 -0.0722 0.0738 -0.98 0.33
## Year1997 -0.0343 0.2048 -0.17 0.87
## Year1998 0.0912 0.2474 0.37 0.71
## Year1999 -0.0542 0.2531 -0.21 0.83
## Year2000 -0.1286 0.1972 -0.65 0.52
## Year2001 -0.0816 0.2220 -0.37 0.71
## Year2002 -0.2741 0.2136 -1.28 0.20
## Year2003 -0.0129 0.2566 -0.05 0.96
## Year2004 -0.0681 0.2333 -0.29 0.77
## Year2005 -0.2579 0.1952 -1.32 0.19
## Year2006 -0.0719 0.2184 -0.33 0.74
```

```

## Year2007          -0.1805      0.1937   -0.93      0.35
## Year2008          -0.3170      0.1965   -1.61      0.11
## Year2009          -0.1754      0.2144   -0.82      0.41
## Year2010          -0.0585      0.2200   -0.27      0.79
## Year2011          -0.0523      0.1929   -0.27      0.79
## Year2012           0.0450      0.2011    0.22      0.82
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.0848, Adjusted R-squared:  -0.00161
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 183 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.378  0.883   0.953   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      5.05e-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 1.605 1      1.267
## Year      1.605 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1454 -0.2595 -0.0432  0.3202  1.2176

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.20839    0.16724    7.23 1.4e-11 ***
## LastAuthorFemale1 -0.00110    0.09183   -0.01    0.99
## Year1997        -0.03916    0.19654   -0.20    0.84
## Year1998         0.09913    0.24299    0.41    0.68
## Year1999        -0.04553    0.25235   -0.18    0.86
## Year2000        -0.12034    0.19127   -0.63    0.53
## Year2001        -0.08607    0.21825   -0.39    0.69
## Year2002        -0.27292    0.21061   -1.30    0.20
## Year2003        -0.00406    0.25231   -0.02    0.99
## Year2004        -0.08430    0.23602   -0.36    0.72
## Year2005        -0.25377    0.18816   -1.35    0.18
## Year2006        -0.07944    0.21050   -0.38    0.71
## Year2007        -0.18176    0.18626   -0.98    0.33
## Year2008        -0.33410    0.18959   -1.76    0.08 .
## Year2009        -0.18633    0.20736   -0.90    0.37
## Year2010        -0.09721    0.20643   -0.47    0.64
## Year2011        -0.06303    0.18877   -0.33    0.74
## Year2012         0.03289    0.19602    0.17    0.87
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.0798, Adjusted R-squared:  -0.00715
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 183 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.372  0.891   0.951   0.917   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      5.05e-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: 198"
## [1] ""

```

```

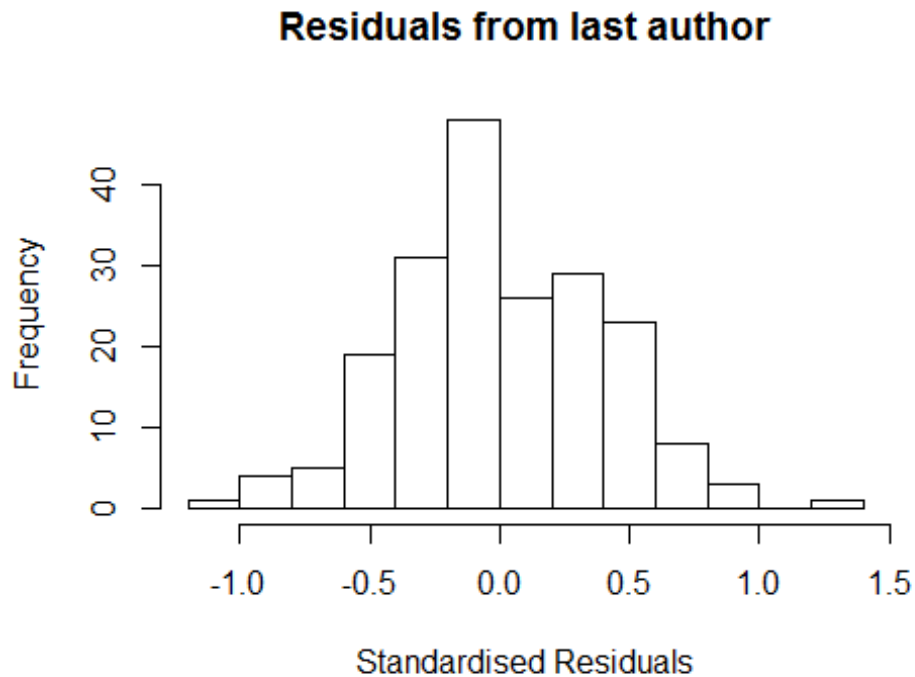
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1200"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    2    1    2    2    1    1    4    1    4    4    2    6    9    4
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    1    0    1    2    1    1    3    1    3    3    1    5    7    4
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    1    0    1    2    1    1    3    1    3    3    1    4    6    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1"
## [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 = 0.6
## 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: 1"

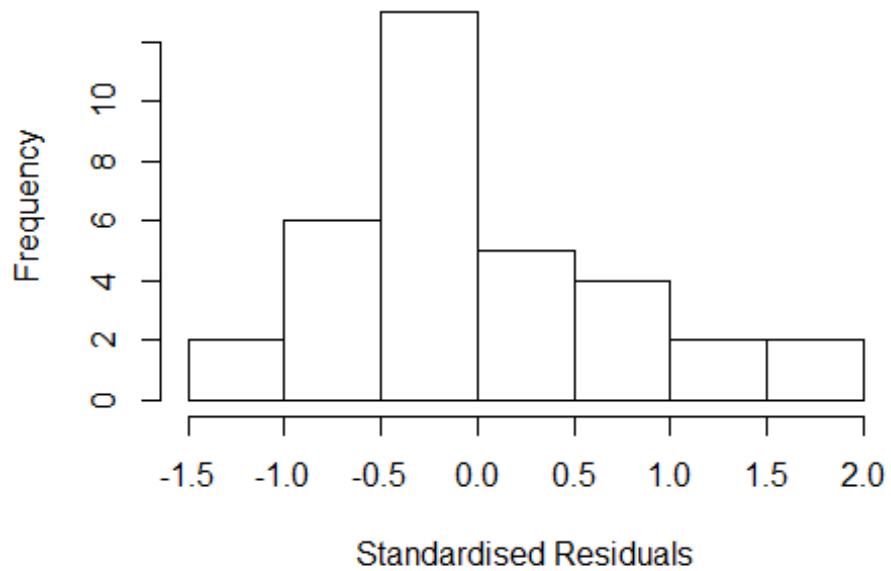
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

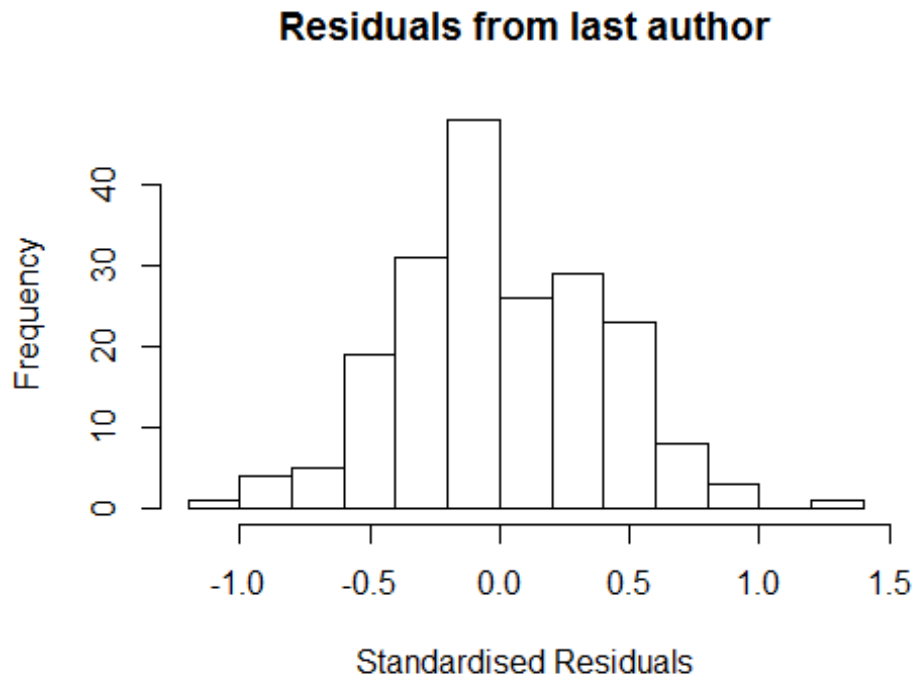


```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
##      Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
##      factors"
## [1] "Regression 3: First author gender, Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.525e+14  1      2.127e+07
## Year              4.525e+14 13      3.662e+00
```

Residuals from first author

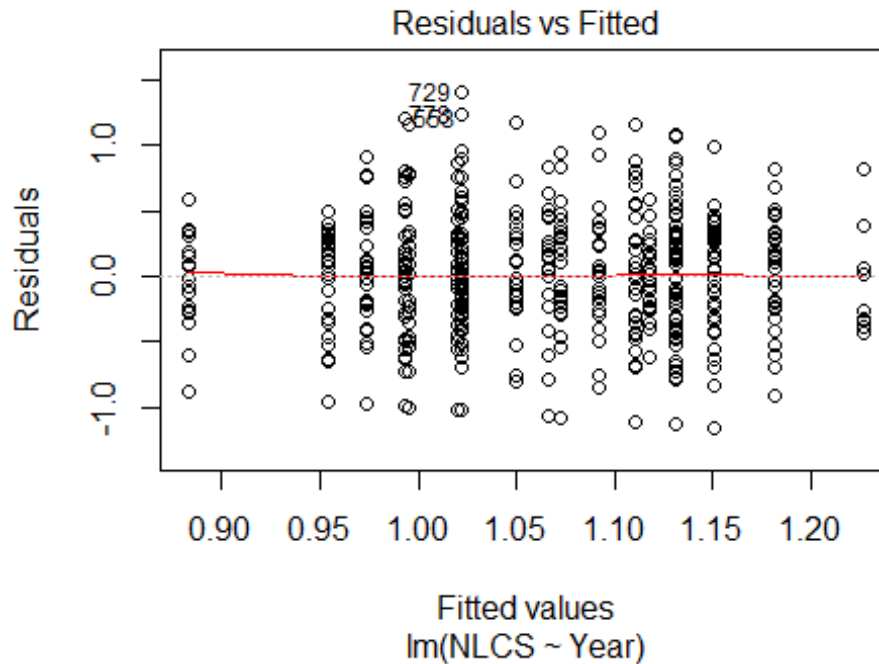


```
## [1] "Regression 4: Last author gender, Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 3.089e+15  1      5.558e+07
## Year              3.089e+15 13      3.942e+00
```



```
## [1] "Sample size for the above analysis: 34"
## [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
## 14 27 29 33 42 37 28 34 25 39 28 30 38 47 49
## 2011 2012
## 68 77
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 25 20 25 30 21 24 29 21 31 25 25 26 36 41
## 2011 2012
## 59 61
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 23 16 25 26 19 23 27 20 26 23 23 24 31 38
## 2011 2012
## 56 59
## [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
```



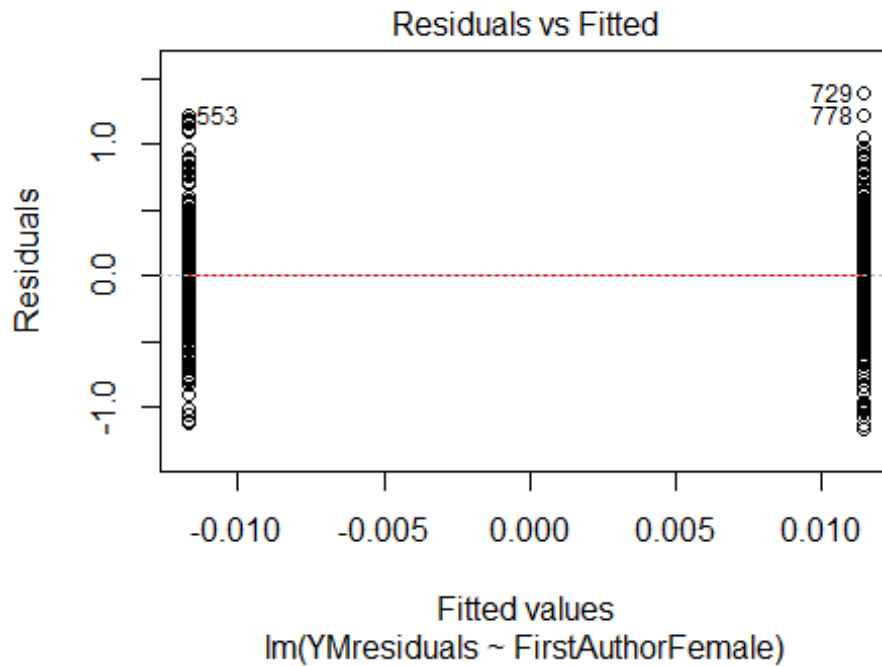
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.54, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 1.6979447322171"
## [1] "Male first author team size 2018 geometric mean: 2.22188951454767"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

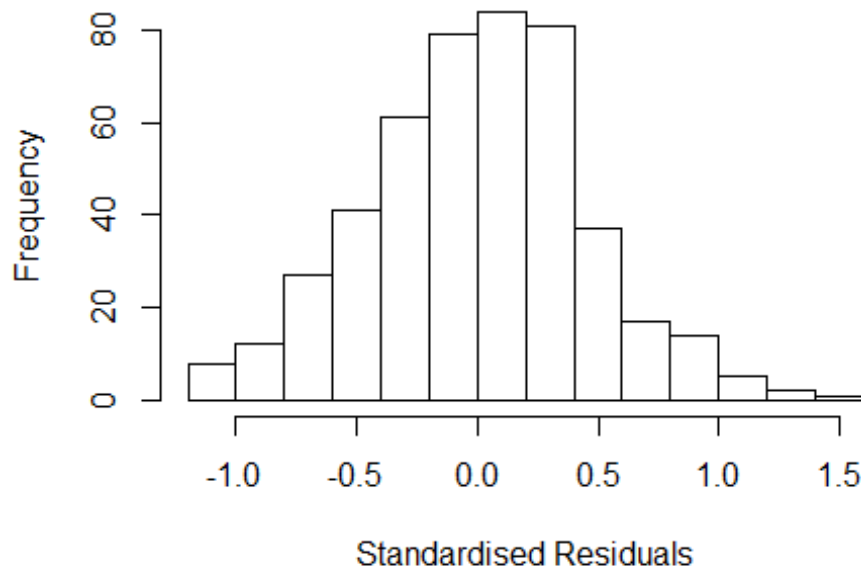
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.87811668297319"
## [1] "Male last author team size 2018 geometric mean: 1.93148120528741"

## 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 = 210, 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 2.091 1 1.446
## LastAuthorFemale 2.077 1 1.441
## UniqueAuthors 1.939 4 1.086
## Year 2.282 16 1.026
```

Residuals from first and last author and team size



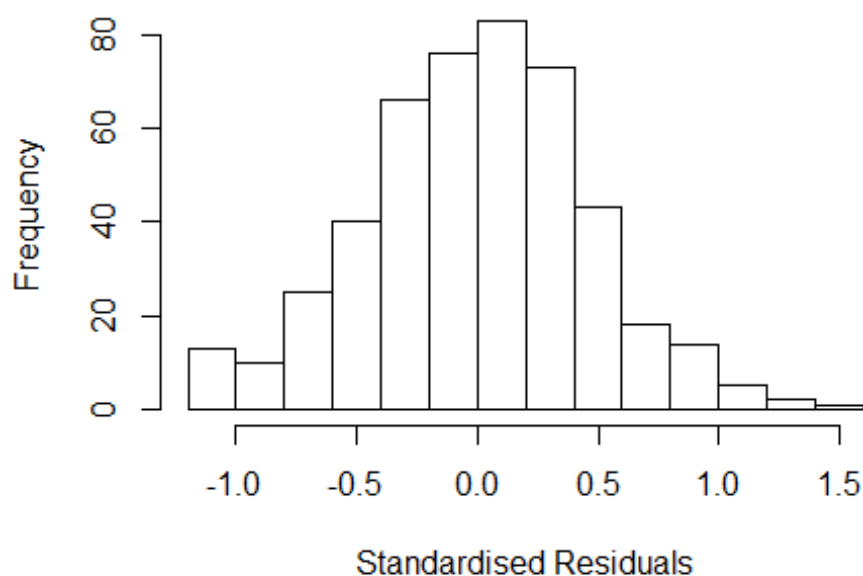
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1817 -0.2812  0.0124  0.2794  1.4189
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.04558    0.11756   8.89  <2e-16 ***
## FirstAuthorFemale1  0.01191    0.06042   0.20   0.8439
## LastAuthorFemale1  0.04934    0.06064   0.81   0.4163
## UniqueAuthors2    0.08905    0.05163   1.72   0.0853 .
## UniqueAuthors3    0.08190    0.07153   1.14   0.2528
## UniqueAuthors4    0.27080    0.09406   2.88   0.0042 **
## UniqueAuthors5    0.28612    0.09759   2.93   0.0035 **
## Year1997          0.00711    0.15835   0.04   0.9642
## Year1998         -0.24235    0.17741  -1.37   0.1726
## Year1999         -0.05329    0.14312  -0.37   0.7098
```

```

## Year2000      -0.09625    0.16839   -0.57    0.5679
## Year2001      -0.10019    0.16122   -0.62    0.5346
## Year2002      -0.08722    0.15596   -0.56    0.5763
## Year2003      -0.14072    0.13980   -1.01    0.3147
## Year2004      -0.23748    0.13722   -1.73    0.0842
## Year2005       0.04187    0.14534    0.29    0.7734
## Year2006      -0.01377    0.13064   -0.11    0.9161
## Year2007      -0.04155    0.14109   -0.29    0.7685
## Year2008      -0.15756    0.15249   -1.03    0.3020
## Year2009      -0.14946    0.13975   -1.07    0.2854
## Year2010      -0.01542    0.14330   -0.11    0.9144
## Year2011      -0.01422    0.13873   -0.10    0.9184
## Year2012      -0.14346    0.13187   -1.09    0.2773
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.432
## Multiple R-squared:  0.0624, Adjusted R-squared:  0.0161
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 45 weights are ~= 1. The remaining 424 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.259  0.860  0.952  0.897  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          2.13e-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 2.128 1          1.459
## LastAuthorFemale 2.118 1          1.455
## Year      1.194 16          1.006

```

Residuals from first and last author



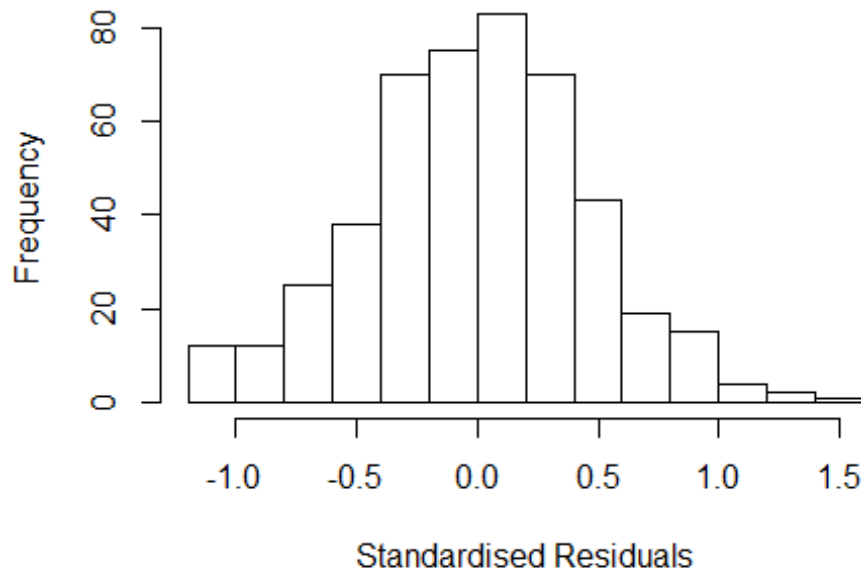
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1687 -0.2829 0.0113 0.3021 1.4382
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.07786 0.12134 8.88 <2e-16 ***
## FirstAuthorFemale1 0.01407 0.06194 0.23 0.82
## LastAuthorFemale1 0.04010 0.06218 0.64 0.52
## Year1997 0.04535 0.16290 0.28 0.78
## Year1998 -0.21050 0.17691 -1.19 0.23
## Year1999 -0.03159 0.14815 -0.21 0.83
## Year2000 -0.06597 0.16794 -0.39 0.69
## Year2001 -0.03350 0.16771 -0.20 0.84
## Year2002 -0.05604 0.15526 -0.36 0.72
## Year2003 -0.14317 0.14391 -0.99 0.32
## Year2004 -0.22423 0.14062 -1.59 0.11
## Year2005 0.05163 0.14807 0.35 0.73
```

```

## Year2006          0.00912    0.13492    0.07    0.95
## Year2007          -0.01940    0.14314   -0.14    0.89
## Year2008          -0.11837    0.15761   -0.75    0.45
## Year2009          -0.11931    0.14297   -0.83    0.40
## Year2010           0.03672    0.14779    0.25    0.80
## Year2011           0.03289    0.14335    0.23    0.82
## Year2012          -0.10809    0.13601   -0.79    0.43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.434
## Multiple R-squared:  0.0331, Adjusted R-squared:  -0.00553
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 436 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.249  0.869  0.953  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      2.13e-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 1.084 1      1.041
## Year              1.084 16      1.003

```

Residuals from first author



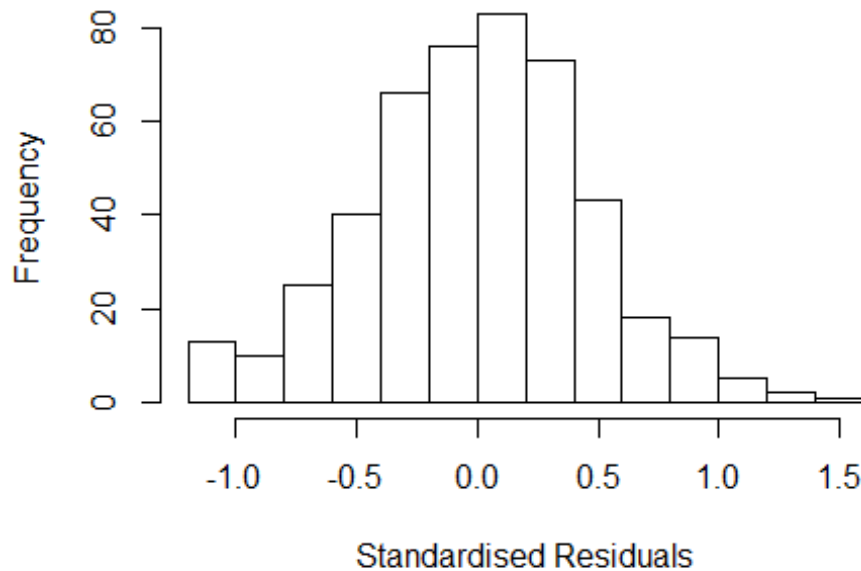
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1589 -0.2781 0.0101 0.3030 1.4029
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08424 0.12248 8.85 <2e-16 ***
## FirstAuthorFemale1 0.03970 0.04444 0.89 0.37
## Year1997 0.03965 0.16334 0.24 0.81
## Year1998 -0.21410 0.17854 -1.20 0.23
## Year1999 -0.03161 0.14894 -0.21 0.83
## Year2000 -0.06016 0.16734 -0.36 0.72
## Year2001 -0.03627 0.16790 -0.22 0.83
## Year2002 -0.05583 0.15554 -0.36 0.72
## Year2003 -0.14190 0.14543 -0.98 0.33
## Year2004 -0.22669 0.14103 -1.61 0.11
## Year2005 0.04899 0.14947 0.33 0.74
## Year2006 0.00811 0.13517 0.06 0.95
```

```

## Year2007          -0.02200    0.14374   -0.15    0.88
## Year2008          -0.11626    0.15872   -0.73    0.46
## Year2009          -0.11944    0.14359   -0.83    0.41
## Year2010           0.03493    0.14724    0.24    0.81
## Year2011           0.03078    0.14376    0.21    0.83
## Year2012          -0.10485    0.13609   -0.77    0.44
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.432
## Multiple R-squared:  0.032, Adjusted R-squared:  -0.00453
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 435 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.270  0.868  0.952  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      2.13e-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 1.074 1          1.036
## Year              1.074 16          1.002

```

Residuals from last author



```
## [1] "List of 0 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
## -1.1682 -0.2809 0.0109 0.3037 1.4499
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08130 0.11989 9.02 <2e-16 ***
## LastAuthorFemale1 0.04883 0.04455 1.10 0.27
## Year1997 0.04668 0.16270 0.29 0.77
## Year1998 -0.21285 0.17530 -1.21 0.23
## Year1999 -0.03306 0.14717 -0.22 0.82
## Year2000 -0.06909 0.16658 -0.41 0.68
## Year2001 -0.03216 0.16746 -0.19 0.85
## Year2002 -0.05775 0.15447 -0.37 0.71
## Year2003 -0.14524 0.14240 -1.02 0.31
## Year2004 -0.22376 0.14036 -1.59 0.11
## Year2005 0.05070 0.14745 0.34 0.73
## Year2006 0.00788 0.13435 0.06 0.95
```

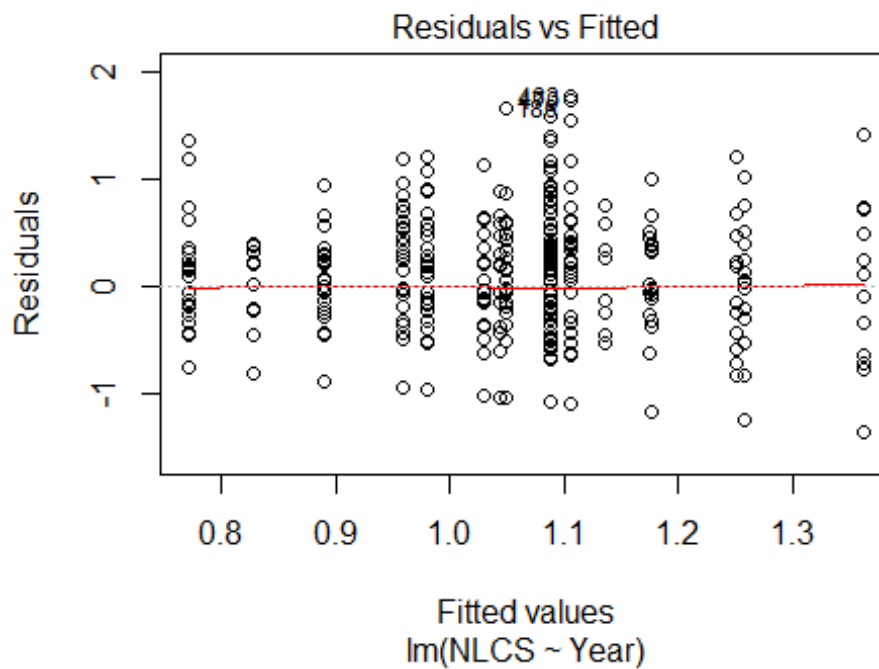


```

## Year2007          -0.01944      0.14262    -0.14      0.89
## Year2008          -0.11911      0.15692    -0.76      0.45
## Year2009          -0.11844      0.14251    -0.83      0.41
## Year2010           0.03806      0.14768      0.26      0.80
## Year2011           0.03292      0.14277      0.23      0.82
## Year2012          -0.10920      0.13518    -0.81      0.42
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.434
## Multiple R-squared:  0.0331, Adjusted R-squared:  -0.00333
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 435 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.242  0.868  0.954  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      2.13e-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: 469"
## [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
##   16   14   21   12   19   11   15   37   18   28   36   38   46   28   45
## 2011 2012
##   46   54
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   14   11   17    9   12    6   14   30   13   24   33   34   36   23   31
## 2011 2012

```

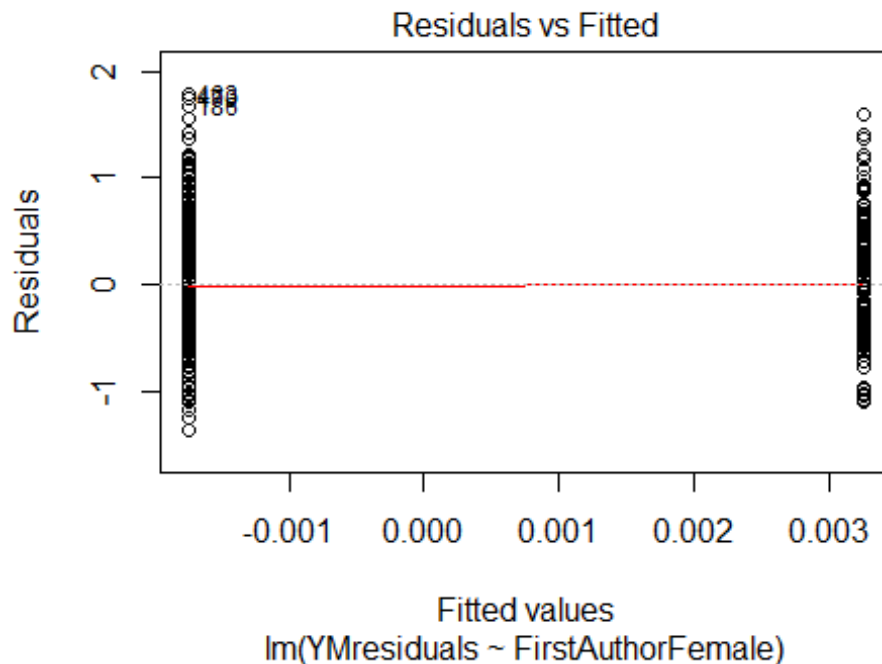
```
## 39 46
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 11 16 9 12 5 14 30 13 24 33 34 36 22 31
## 2011 2012
## 38 44
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 18, df = 16, p-value = 0.3
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.054, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 1.39348431091769"
## [1] "Male first author team size 2018 geometric mean: 1.02122660631536"
## 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 = 7e-04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.26327525472446"
## [1] "Male last author team size 2018 geometric mean: 1.12253056625947"

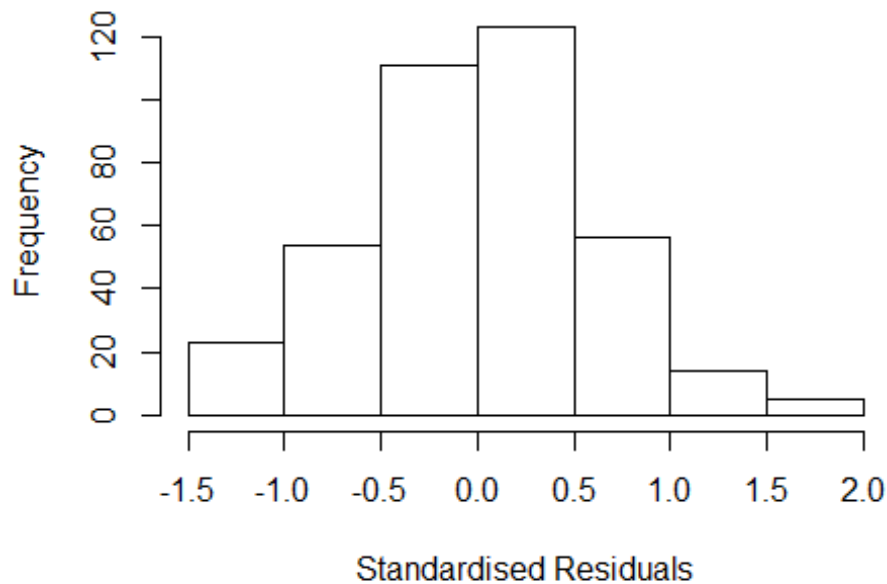
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 480, 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	10.37	1	3.221
LastAuthorFemale	10.34	1	3.215
UniqueAuthors	39.62	4	1.584
Year	44.83	16	1.126

Residuals from first and last author and team size



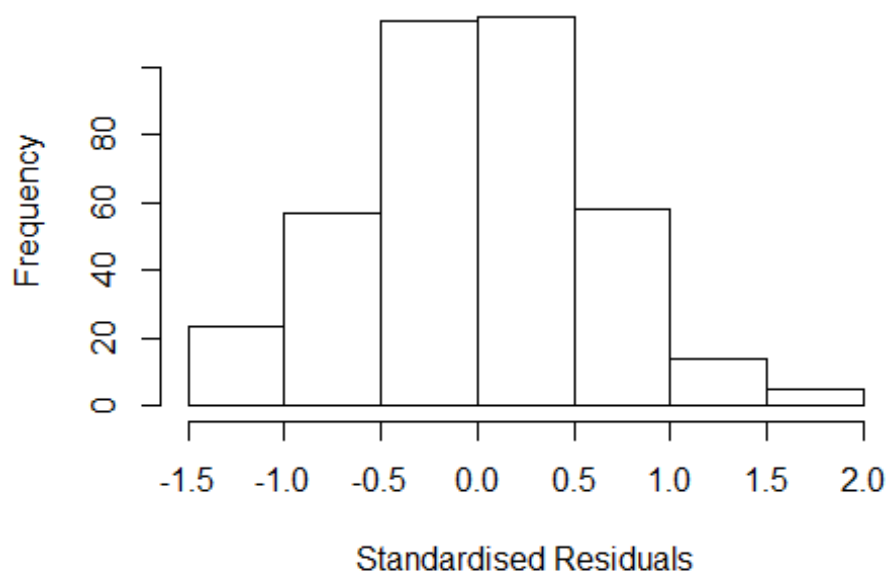
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27623 -0.44500  0.00966  0.39637  1.85924
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2535    0.1624   7.72 1.1e-13 ***
## FirstAuthorFemale1 -0.1393    0.1741  -0.80  0.4241
## LastAuthorFemale1  0.1788    0.1738   1.03  0.3043
## UniqueAuthors2     0.1588    0.1123   1.41  0.1582
## UniqueAuthors3     0.4105    0.2074   1.98  0.0486 *
## UniqueAuthors4     0.1890    0.9361   0.20  0.8401
## UniqueAuthors5     0.7012    0.1486   4.72 3.4e-06 ***
## Year1997          -0.4268    0.1970  -2.17  0.0309 *
## Year1998          -0.1403    0.2083  -0.67  0.5012
## Year1999          -0.1287    0.2325  -0.55  0.5802
```

```

## Year2000          -0.0353      0.2417   -0.15   0.8839
## Year2001          -0.0676      0.2595   -0.26   0.7947
## Year2002          -0.2308      0.2240   -1.03   0.3034
## Year2003          -0.5340      0.1928   -2.77   0.0059 **
## Year2004           0.0227      0.2780    0.08   0.9350
## Year2005          -0.2678      0.2040   -1.31   0.1900
## Year2006          -0.2267      0.2022   -1.12   0.2628
## Year2007          -0.3370      0.1996   -1.69   0.0922 .
## Year2008          -0.3227      0.1937   -1.67   0.0966 .
## Year2009          -0.2784      0.2018   -1.38   0.1684
## Year2010          -0.3810      0.1823   -2.09   0.0373 *
## Year2011          -0.2298      0.2203   -1.04   0.2976
## Year2012          -0.2139      0.1950   -1.10   0.2735
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.634
## Multiple R-squared:  0.0649, Adjusted R-squared:  0.00823
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 350 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.370  0.879  0.953  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          2.59e-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 7.044 1      2.654
## LastAuthorFemale  7.089 1      2.663
## Year              1.254 16      1.007

```

Residuals from first and last author



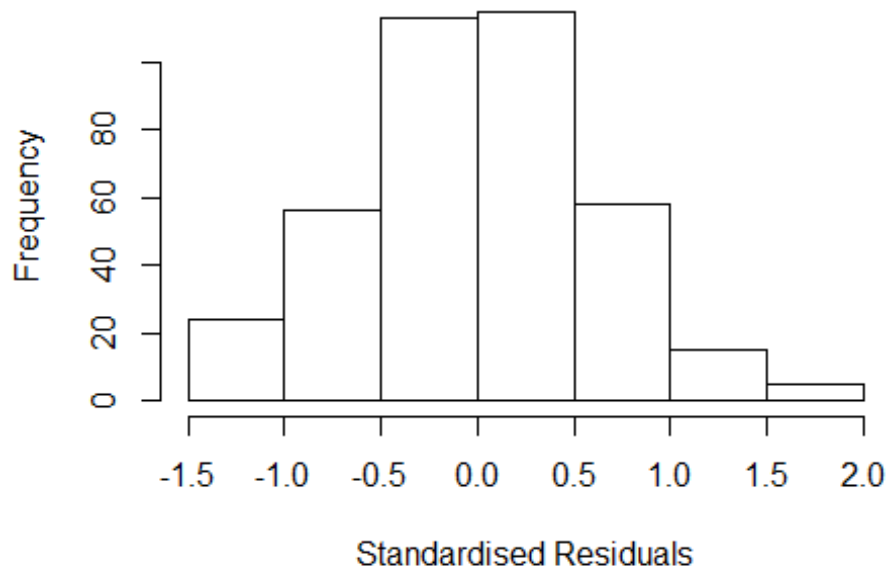
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.33199 -0.43045 -0.00392 0.39945 1.81760
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2651 0.1609 7.86 4.3e-14 ***
## FirstAuthorFemale1 -0.1421 0.1808 -0.79 0.4323
## LastAuthorFemale1 0.1924 0.1803 1.07 0.2867
## Year1997 -0.4416 0.1956 -2.26 0.0245 *
## Year1998 -0.1448 0.2068 -0.70 0.4842
## Year1999 -0.1412 0.2309 -0.61 0.5413
## Year2000 -0.0337 0.2441 -0.14 0.8902
## Year2001 -0.0814 0.2574 -0.32 0.7520
## Year2002 -0.2326 0.2218 -1.05 0.2950
## Year2003 -0.5356 0.1929 -2.78 0.0058 **
## Year2004 0.0669 0.2795 0.24 0.8110
## Year2005 -0.2511 0.2019 -1.24 0.2144
```

```

## Year2006          -0.2249      0.2001   -1.12   0.2617
## Year2007          -0.3007      0.2033   -1.48   0.1400
## Year2008          -0.3226      0.1944   -1.66   0.0979 .
## Year2009          -0.2742      0.1984   -1.38   0.1677
## Year2010          -0.3796      0.1808   -2.10   0.0365 *
## Year2011          -0.1997      0.2162   -0.92   0.3563
## Year2012          -0.1834      0.1900   -0.97   0.3350
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.638
## Multiple R-squared:  0.0495, Adjusted R-squared:  0.00292
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 355 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.397  0.884  0.955  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      2.59e-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 1.1 1      1.049
## Year              1.1 16      1.003

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.356299 -0.439244 0.000267 0.389243 1.820165
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2693 0.1611 7.88 3.8e-14 ***
## FirstAuthorFemale1 0.0329 0.0716 0.46 0.6462
## Year1997 -0.4408 0.1959 -2.25 0.0251 *
## Year1998 -0.1415 0.2074 -0.68 0.4953
## Year1999 -0.1439 0.2319 -0.62 0.5353
## Year2000 -0.0505 0.2399 -0.21 0.8332
## Year2001 -0.0821 0.2592 -0.32 0.7517
## Year2002 -0.2305 0.2229 -1.03 0.3016
## Year2003 -0.5278 0.1933 -2.73 0.0066 **
## Year2004 0.0870 0.2757 0.32 0.7526
## Year2005 -0.2501 0.2018 -1.24 0.2161
## Year2006 -0.2158 0.1993 -1.08 0.2796
```

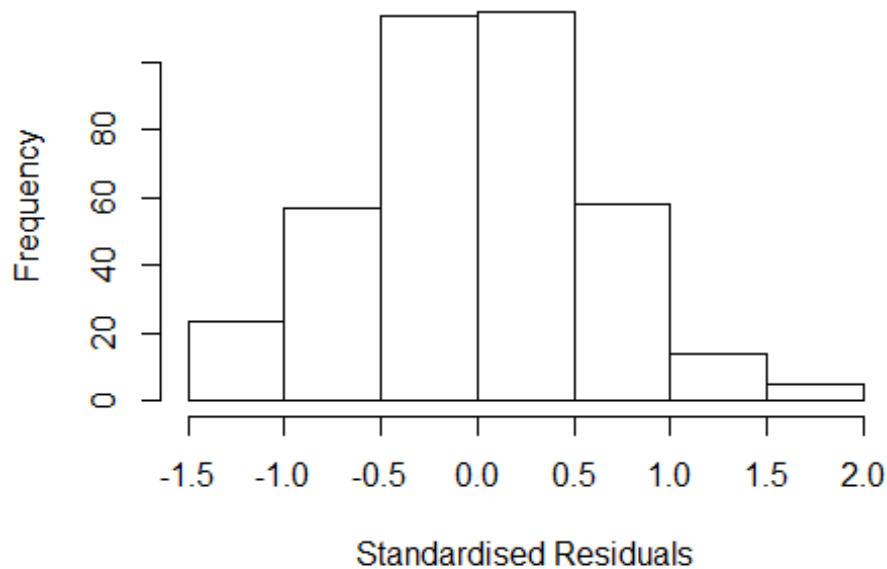


```

## Year2007          -0.2945      0.2033   -1.45    0.1483
## Year2008          -0.3195      0.1947   -1.64    0.1016
## Year2009          -0.2719      0.1982   -1.37    0.1708
## Year2010          -0.3856      0.1806   -2.13    0.0335 *
## Year2011          -0.2065      0.2151   -0.96    0.3377
## Year2012          -0.1827      0.1929   -0.95    0.3440
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.636
## Multiple R-squared:  0.0461, Adjusted R-squared:  0.00199
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 356 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.393  0.880   0.954   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      2.59e-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 1.12 1          1.058
## Year              1.12 16          1.004

```

Residuals from last author



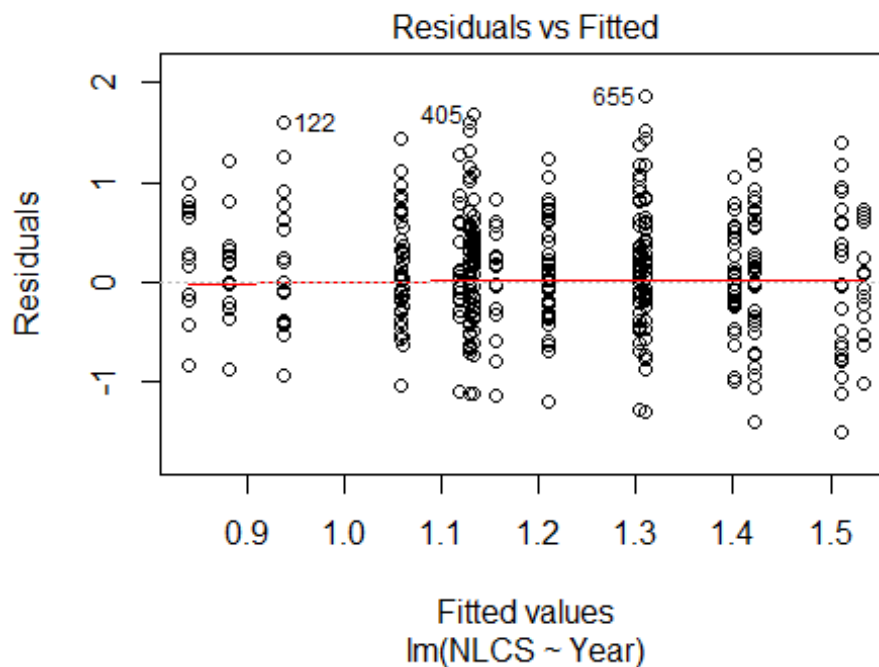
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.34379 -0.43698 -0.00154  0.39014  1.82654
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2621     0.1608   7.85 4.6e-14 ***
## LastAuthorFemale1  0.0634     0.0718   0.88  0.3779
## Year1997         -0.4423     0.1953  -2.26  0.0241 *
## Year1998         -0.1472     0.2063  -0.71  0.4760
## Year1999         -0.1393     0.2302  -0.61  0.5455
## Year2000         -0.0484     0.2398  -0.20  0.8403
## Year2001         -0.0810     0.2563  -0.32  0.7520
## Year2002         -0.2343     0.2212  -1.06  0.2901
## Year2003         -0.5309     0.1931  -2.75  0.0063 **
## Year2004          0.0817     0.2757   0.30  0.7671
## Year2005         -0.2522     0.2020  -1.25  0.2127
## Year2006         -0.2240     0.1994  -1.12  0.2620
```

```

## Year2007          -0.2981      0.2024    -1.47    0.1416
## Year2008          -0.3250      0.1943    -1.67    0.0952 .
## Year2009          -0.2761      0.1986    -1.39    0.1653
## Year2010          -0.3846      0.1801    -2.14    0.0333 *
## Year2011          -0.2056      0.2154    -0.95    0.3404
## Year2012          -0.1889      0.1914    -0.99    0.3243
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.637
## Multiple R-squared:  0.0476, Adjusted R-squared:  0.00362
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 356 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.392  0.884  0.954  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      2.59e-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: 386"
## [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
##   11   23   21   23   17   26   19   26   32   33   42   49   46   54   39
## 2011 2012
##   72   62
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11   18   18   20   14   21   18   22   23   25   34   42   43   42   29
## 2011 2012

```

```
## 55 49
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 18 18 20 13 19 18 21 22 23 34 41 43 42 29
## 2011 2012
## 52 46
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.5
```



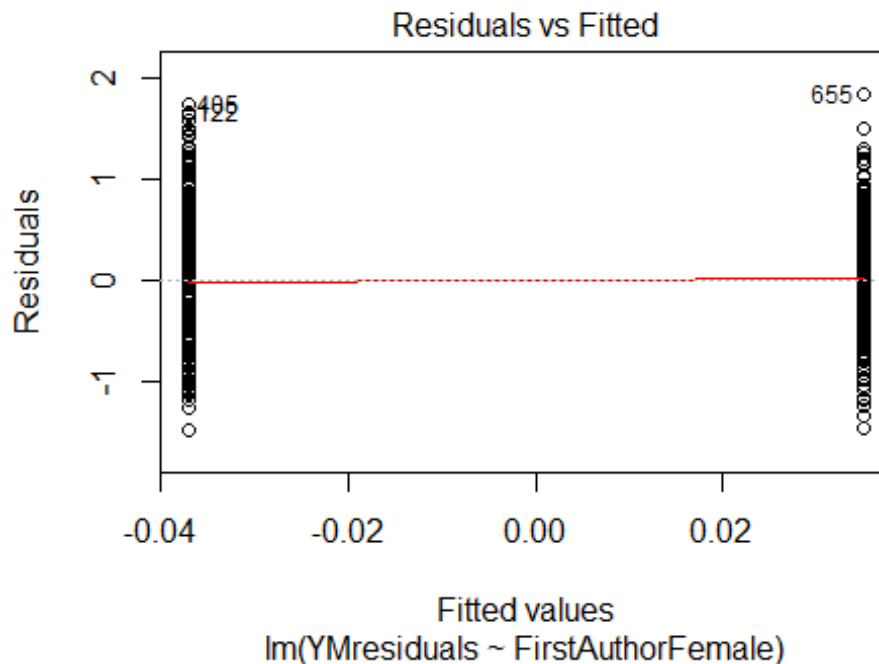
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.4, df = 1, p-value = 0.04

## [1] "Female first author team size 2018 geometric mean: 1.37410881031664"
## [1] "Male first author team size 2018 geometric mean: 1.25573079078343"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

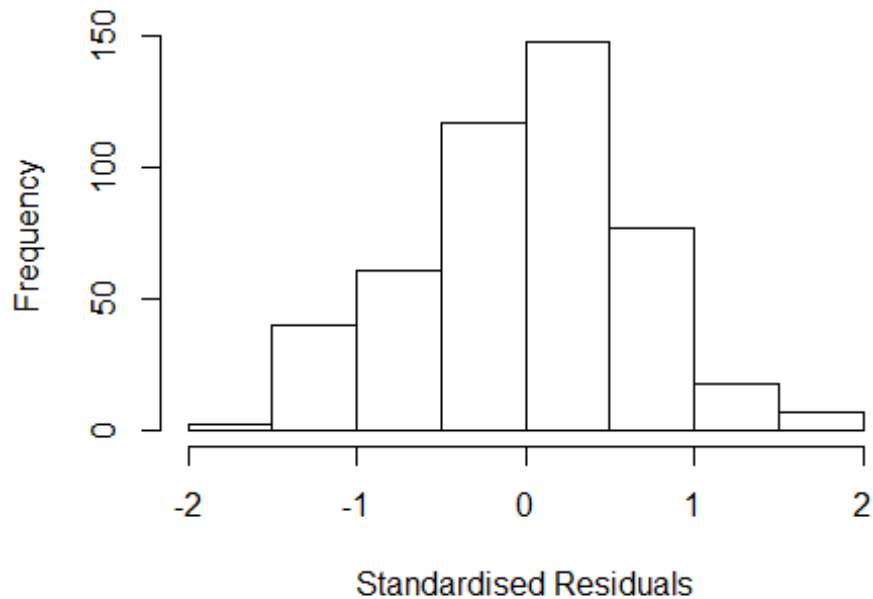
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.51365727747059"
## [1] "Male last author team size 2018 geometric mean: 1.05075663865322"

## 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 = 210, 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 3.425  1      1.851
## LastAuthorFemale  3.133  1      1.770
## UniqueAuthors    2.587  4      1.126
## Year              2.563 16      1.030
```

Residuals from first and last author and team size



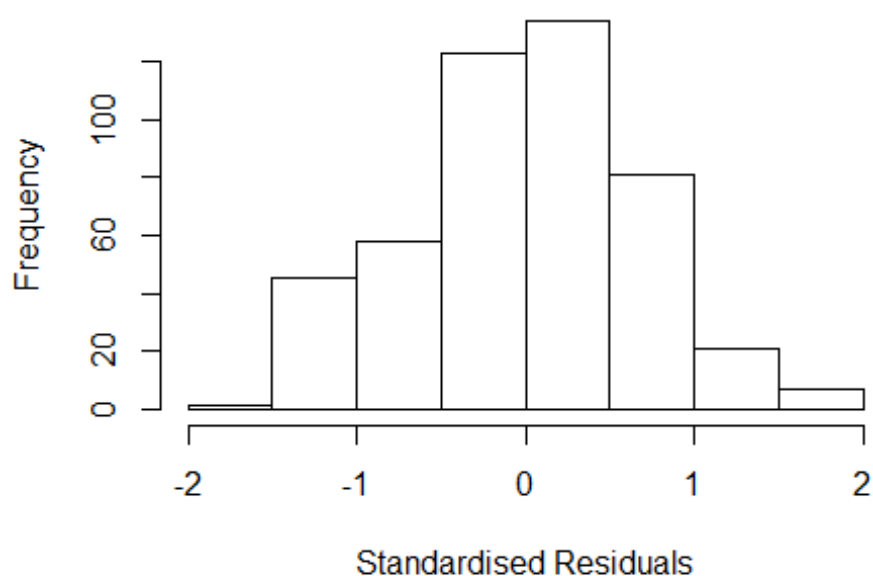
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5529 -0.4142  0.0273  0.4386  1.9945
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9287    0.1145   8.11 4.9e-15 ***
## FirstAuthorFemale1  0.0291    0.1104   0.26  0.79228
## LastAuthorFemale1  0.0192    0.1058   0.18  0.85609
## UniqueAuthors2    0.2645    0.0796   3.32  0.00097 ***
## UniqueAuthors3    0.3467    0.1246   2.78  0.00563 **
## UniqueAuthors4    0.3300    0.1722   1.92  0.05591 .
## UniqueAuthors5   -0.1484    0.1936  -0.77  0.44380
## Year1997          0.1882    0.1864   1.01  0.31323
## Year1998         -0.1389    0.1892  -0.73  0.46309
## Year1999         -0.1176    0.1762  -0.67  0.50470
```

```

## Year2000          0.5735      0.1900      3.02  0.00269 **
## Year2001         -0.1153      0.2019     -0.57  0.56835
## Year2002          0.2761      0.2005      1.38  0.16928
## Year2003          0.1357      0.1740      0.78  0.43576
## Year2004          0.2751      0.1928      1.43  0.15427
## Year2005          0.6005      0.2271      2.64  0.00847 **
## Year2006          0.3748      0.1668      2.25  0.02509 *
## Year2007          0.1278      0.1300      0.98  0.32641
## Year2008          0.1092      0.1382      0.79  0.42977
## Year2009          0.0357      0.1508      0.24  0.81309
## Year2010          0.3610      0.1400      2.58  0.01024 *
## Year2011          0.0895      0.1430      0.63  0.53163
## Year2012          0.2016      0.1553      1.30  0.19492
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.669
## Multiple R-squared:  0.116, Adjusted R-squared:  0.0725
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 43 weights are ~= 1. The remaining 427 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.354  0.862  0.953   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      2.13e-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 3.195 1      1.787
## LastAuthorFemale  3.074 1      1.753
## Year              1.274 16      1.008

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.5810 -0.4108 0.0152 0.4522 1.8431
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9883 0.1243 7.95 1.5e-14 ***
## FirstAuthorFemale1 0.0685 0.1111 0.62 0.5379
## LastAuthorFemale1 0.0259 0.1088 0.24 0.8121
## Year1997 0.1473 0.1866 0.79 0.4301
## Year1998 -0.2101 0.1914 -1.10 0.2730
## Year1999 -0.1576 0.1815 -0.87 0.3857
## Year2000 0.5032 0.1965 2.56 0.0108 *
## Year2001 -0.1824 0.2120 -0.86 0.3902
## Year2002 0.2956 0.2219 1.33 0.1835
## Year2003 0.0849 0.1788 0.47 0.6353
## Year2004 0.2918 0.2029 1.44 0.1511
## Year2005 0.5927 0.2229 2.66 0.0081 **
```

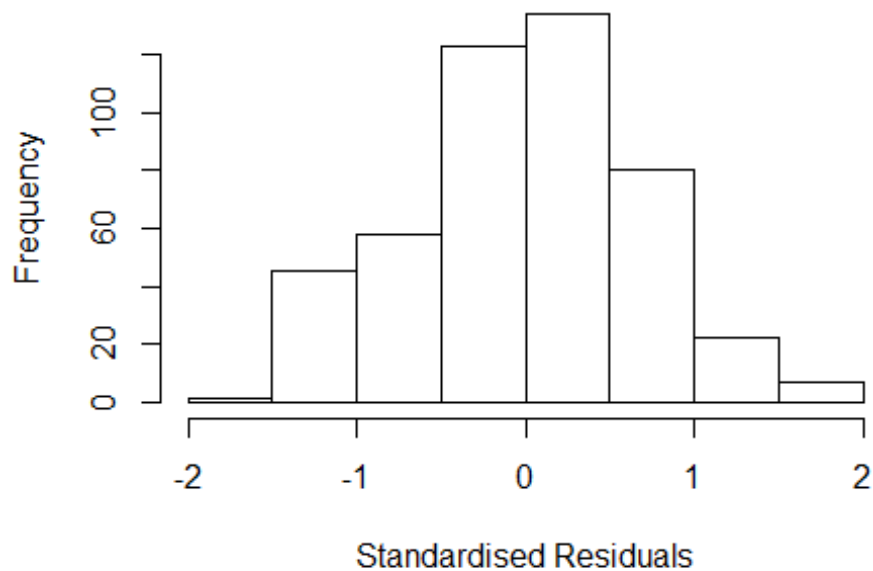


```

## Year2006          0.3940      0.1747      2.26      0.0246 *
## Year2007          0.1429      0.1417      1.01      0.3137
## Year2008          0.1080      0.1486      0.73      0.4676
## Year2009          0.0269      0.1605      0.17      0.8669
## Year2010          0.3714      0.1496      2.48      0.0134 *
## Year2011          0.0932      0.1515      0.62      0.5386
## Year2012          0.2472      0.1593      1.55      0.1214
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.679
## Multiple R-squared:  0.0832, Adjusted R-squared:  0.0466
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 40 weights are ~= 1. The remaining 430 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.442  0.874   0.953   0.911   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      2.13e-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 1.143 1          1.069
## Year              1.143 16          1.004

```

Residuals from first author



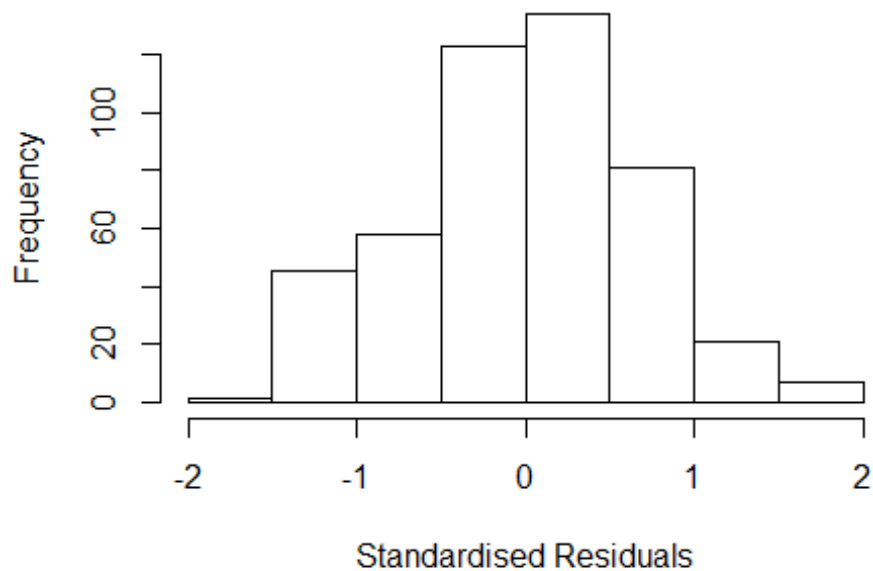
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5844 -0.4098 0.0159 0.4456 1.8468
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9893 0.1236 8.01 1e-14 ***
## FirstAuthorFemale1 0.0903 0.0666 1.36 0.176
## Year1997 0.1487 0.1861 0.80 0.425
## Year1998 -0.2085 0.1910 -1.09 0.275
## Year1999 -0.1557 0.1810 -0.86 0.390
## Year2000 0.5052 0.1956 2.58 0.010 *
## Year2001 -0.1822 0.2117 -0.86 0.390
## Year2002 0.2954 0.2212 1.34 0.183
## Year2003 0.0858 0.1783 0.48 0.630
## Year2004 0.2892 0.1997 1.45 0.148
## Year2005 0.5951 0.2234 2.66 0.008 **
## Year2006 0.3979 0.1750 2.27 0.023 *
```

```

## Year2007          0.1449      0.1414      1.02      0.306
## Year2008          0.1082      0.1480      0.73      0.465
## Year2009          0.0265      0.1599      0.17      0.868
## Year2010          0.3735      0.1488      2.51      0.012 *
## Year2011          0.0951      0.1503      0.63      0.527
## Year2012          0.2466      0.1587      1.55      0.121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.678
## Multiple R-squared:  0.0831, Adjusted R-squared:  0.0486
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 429 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.438  0.873   0.951   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      2.13e-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

```

Residuals from last author



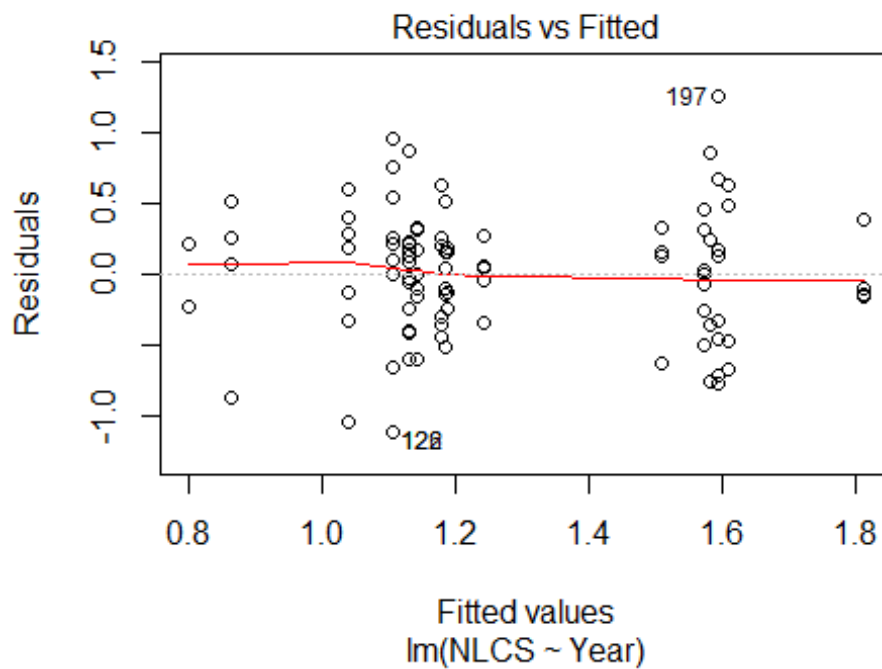
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5798 -0.4118 0.0192 0.4549 1.8441
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0028 0.1235 8.12 4.4e-15 ***
## LastAuthorFemale1 0.0831 0.0654 1.27 0.2045
## Year1997 0.1390 0.1869 0.74 0.4574
## Year1998 -0.2174 0.1925 -1.13 0.2593
## Year1999 -0.1715 0.1812 -0.95 0.3445
## Year2000 0.4970 0.1971 2.52 0.0120 *
## Year2001 -0.1931 0.2126 -0.91 0.3641
## Year2002 0.2911 0.2243 1.30 0.1952
## Year2003 0.0758 0.1792 0.42 0.6726
## Year2004 0.2933 0.2056 1.43 0.1543
## Year2005 0.5771 0.2210 2.61 0.0093 **
## Year2006 0.3800 0.1747 2.18 0.0301 *
```

```

## Year2007          0.1339      0.1424      0.94      0.3477
## Year2008          0.1012      0.1503      0.67      0.5009
## Year2009          0.0221      0.1624      0.14      0.8918
## Year2010          0.3601      0.1496      2.41      0.0165 *
## Year2011          0.0833      0.1521      0.55      0.5842
## Year2012          0.2430      0.1612      1.51      0.1322
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.679
## Multiple R-squared:  0.0826, Adjusted R-squared:  0.0481
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 40 weights are ~= 1. The remaining 430 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.441  0.875   0.952   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      2.13e-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: 470"
## [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
##   10   9   8   6  10  14   4  10   5   6   9  12   7   6   6
## 2011 2012
##    9    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    2    4    5    7    7    4    7    4    5    7   10    4    6    4
## 2011 2012

```

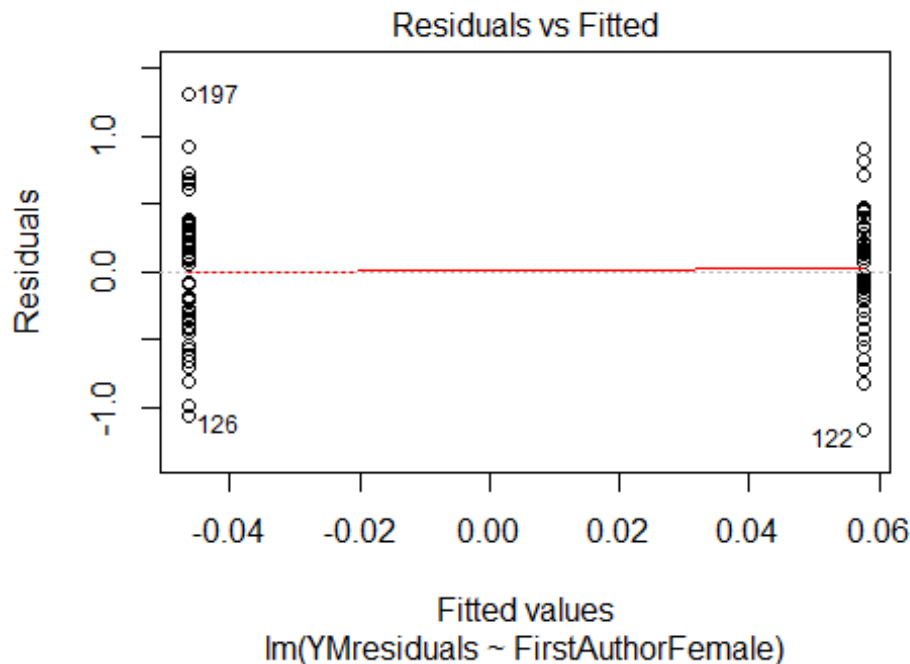
```
##      8      4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6      1      4      5      7      6      4      7      4      5      7      10      4      6      4
## 2011 2012
##      7      4
## [1] "Heteroscedasticity checks, confirming that there are 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 = 0.72, df = 1, p-value = 0.4
## [1] "Female first author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male first author team size 2018 geometric mean: 1.21901365420448"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

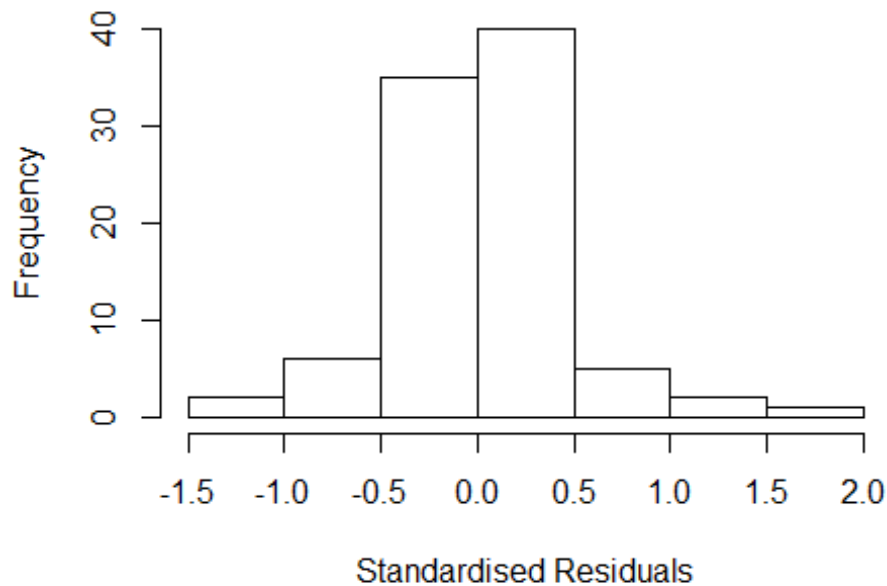
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 13, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 1.48773782616449"

## 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 = 6, 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  6.870  1      2.621
## LastAuthorFemale  6.163  1      2.483
## UniqueAuthors    133.587  4      1.844
## Year              517.414 16      1.216
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.2029 -0.2520 0.0384 0.2266 1.8004
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9335 0.1283 7.27 4.5e-10 ***
## FirstAuthorFemale1 0.0877 0.1155 0.76 0.450
## LastAuthorFemale1 0.0904 0.1371 0.66 0.512
## UniqueAuthors2 0.4788 0.1154 4.15 9.5e-05 ***
## UniqueAuthors3 0.1940 0.1484 1.31 0.196
## UniqueAuthors4 -0.1703 0.4232 -0.40 0.689
## UniqueAuthors5 0.2283 0.2051 1.11 0.270
## Year1997 -0.8383 0.1070 -7.83 4.4e-11 ***
## Year1998 0.4345 0.2537 1.71 0.091 .
## Year1999 0.2166 0.1415 1.53 0.130
```

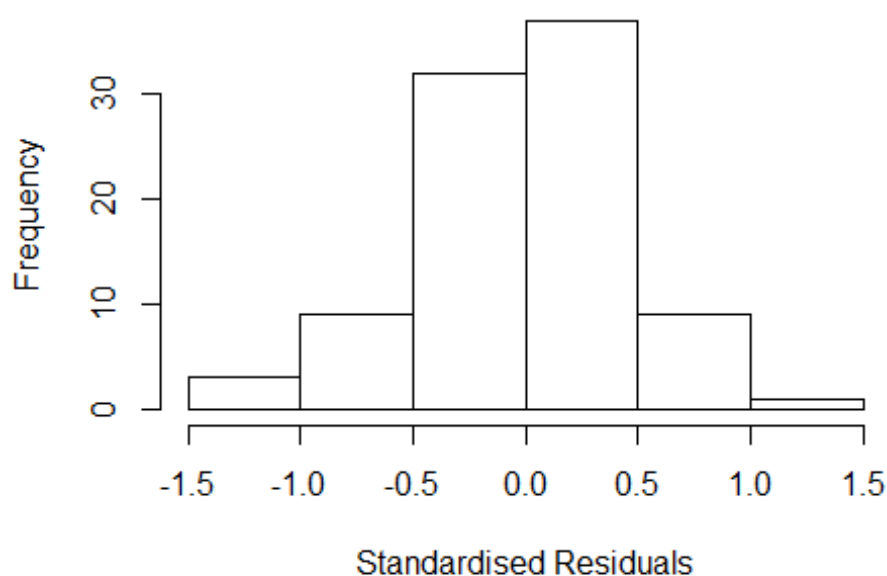


```

## Year2000          -0.1480      0.1315    -1.13      0.264
## Year2001          -0.1292      0.1345    -0.96      0.340
## Year2002          -0.1623      0.2907    -0.56      0.578
## Year2003           0.0823      0.2361     0.35      0.729
## Year2004           0.6795      0.4387     1.55      0.126
## Year2005          -0.0184      0.1783    -0.10      0.918
## Year2006           0.3520      0.1700     2.07      0.042 *
## Year2007           0.0913      0.3139     0.29      0.772
## Year2008           0.5002      0.2234     2.24      0.028 *
## Year2009          -0.0323      0.2582    -0.12      0.901
## Year2010          -0.0115      0.1266    -0.09      0.928
## Year2011           0.2883      0.2868     1.01      0.318
## Year2012           0.5439      0.3617     1.50      0.137
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.369, Adjusted R-squared:  0.164
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 85 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0118 0.8910 0.9610 0.9020 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          1.10e-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 7.485 1          2.736
## LastAuthorFemale  5.201 1          2.280
## Year             11.625 16          1.080

```

Residuals from first and last author



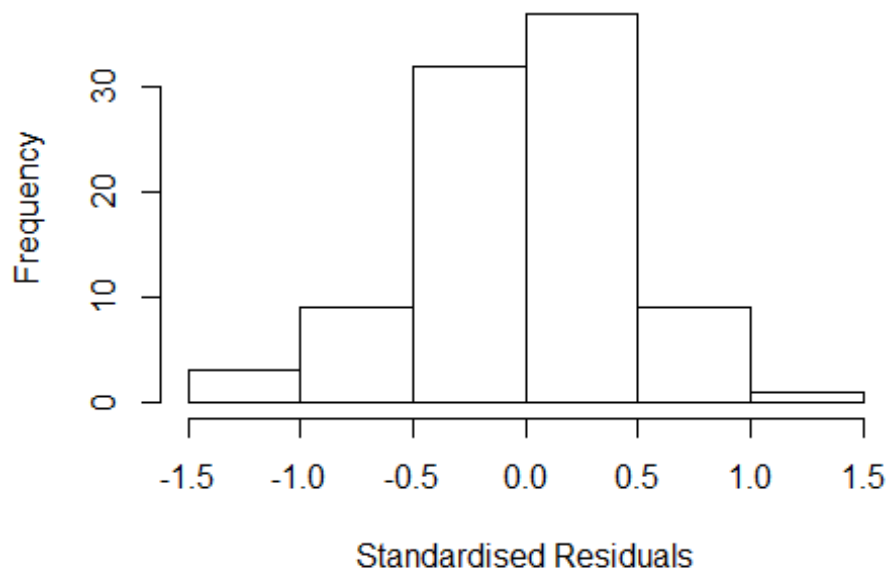
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.224 -0.285 0.039 0.276 1.481
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1564 0.1317 8.78 5.4e-13 ***
## FirstAuthorFemale1 0.1542 0.1424 1.08 0.28249
## LastAuthorFemale1 -0.0596 0.1490 -0.40 0.69017
## Year1997 -0.5824 0.1317 -4.42 3.4e-05 ***
## Year1998 0.3326 0.2354 1.41 0.16192
## Year1999 0.0355 0.1546 0.23 0.81916
## Year2000 -0.0655 0.1732 -0.38 0.70670
## Year2001 -0.1057 0.1871 -0.56 0.57389
## Year2002 -0.2961 0.2877 -1.03 0.30683
## Year2003 -0.1039 0.2434 -0.43 0.67080
## Year2004 0.4226 0.4574 0.92 0.35868
## Year2005 -0.0876 0.2708 -0.32 0.74727
```

```

## Year2006          0.3646      0.1710      2.13  0.03635 *
## Year2007         -0.0265      0.3213     -0.08  0.93449
## Year2008          0.5786      0.1678      3.45  0.00095 ***
## Year2009         -0.0661      0.2490     -0.27  0.79152
## Year2010         -0.0156      0.1612     -0.10  0.92297
## Year2011          0.2143      0.3953      0.54  0.58943
## Year2012          0.3867      0.4173      0.93  0.35715
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.224, Adjusted R-squared:  0.0301
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.252  0.887  0.962  0.906  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      1.10e-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 5.088 1      2.256
## Year              5.088 16      1.052

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2686 -0.2926  0.0544  0.2763  1.4893
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.14590    0.12617   9.08 1.3e-13 ***
## FirstAuthorFemale1 0.12156    0.11717   1.04  0.3029
## Year1997      -0.57190    0.12617  -4.53 2.2e-05 ***
## Year1998       0.35386    0.22491   1.57  0.1200
## Year1999       0.04736    0.14745   0.32  0.7490
## Year2000      -0.06035    0.17743  -0.34  0.7347
## Year2001      -0.08202    0.17316  -0.47  0.6371
## Year2002      -0.28222    0.29212  -0.97  0.3372
## Year2003      -0.10608    0.24486  -0.43  0.6661
## Year2004       0.40781    0.45053   0.91  0.3683
## Year2005      -0.08057    0.26871  -0.30  0.7651
## Year2006       0.36033    0.17424   2.07  0.0422 *
```

```

## Year2007          0.00117    0.31495    0.00    0.9970
## Year2008          0.56902    0.16986    3.35    0.0013 **
## Year2009         -0.06614    0.25270   -0.26    0.7943
## Year2010         -0.01742    0.16144   -0.11    0.9144
## Year2011          0.21677    0.39087    0.55    0.5809
## Year2012          0.39367    0.41066    0.96    0.3409
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.224, Adjusted R-squared:  0.0436
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.224  0.888  0.960  0.903  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      1.10e-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 3.577 1      1.891
## Year              3.577 16      1.041

## [1] "List of 0 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
## -1.255 -0.276  0.023  0.263  1.469

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.171000   0.143505   8.16 7.0e-12 ***
## LastAuthorFemale1 0.035333   0.119597   0.30 0.76850
## Year1997      -0.597000   0.143505  -4.16 8.6e-05 ***
## Year1998       0.365115   0.239654   1.52 0.13195
## Year1999       0.064346   0.159737   0.40 0.68825
## Year2000      -0.033387   0.176700  -0.19 0.85066
## Year2001      -0.026525   0.175087  -0.15 0.88000
## Year2002      -0.244256   0.312855  -0.78 0.43748
## Year2003      -0.081471   0.244516  -0.33 0.73995
## Year2004       0.419953   0.450060   0.93 0.35384
## Year2005      -0.105513   0.276599  -0.38 0.70397
## Year2006       0.382620   0.173556   2.20 0.03064 *
## Year2007       0.049070   0.323288   0.15 0.87978
## Year2008       0.607308   0.170414   3.56 0.00065 ***
## Year2009      -0.021763   0.225816  -0.10 0.92349
## Year2010      -0.000341   0.160825   0.00 0.99831
## Year2011       0.211991   0.375514   0.56 0.57412
## Year2012       0.374715   0.451163   0.83 0.40893
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.211, Adjusted R-squared:  0.027
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 84 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.239  0.878  0.955  0.899  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      1.10e-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: 91"
## [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]"
##
## 1998 2000 2001 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    2    1    1    2    3    3    2    3    1    1    5    1
##
## 1998 2000 2001 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    2    1    1    2    3    3    1    2    1    1    4    1
##
## 1998 2000 2001 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    2    1    1    2    3    3    1    2    1    1    4    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1"
## [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: 22"
## [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]"
##
## 2006 2008 2009 2012
##    1    3    5    2
##
## 2006 2008 2009 2012
##    1    1    5    1
##
## 2006 2008 2009 2012
##    0    1    5    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.37972966146121"
## [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"

```

```

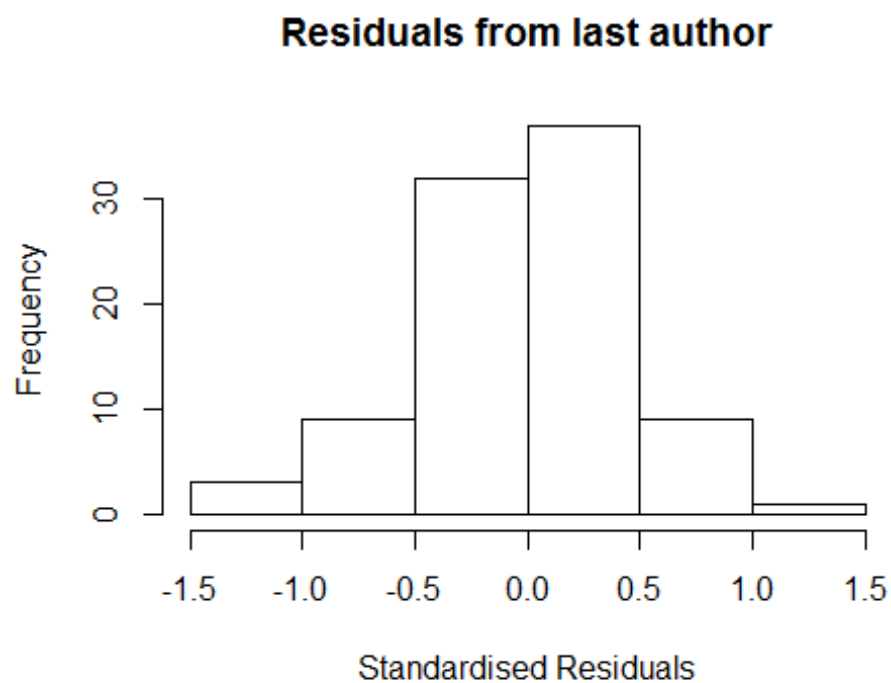
## [1] "Regression 3: First author gender, Year as factors"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 7"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1207"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      8      6      4      6     12      8      1      4      6     11     15      9     18     20     16
## 2011 2012
##     29     23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      7      6      2      6      9      7      1      3      4     10     14      7     18     17     10
## 2011 2012
##     29     21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      7      6      2      6      9      7      1      3      3     10     13      7     18     15     10
## 2011 2012
##     27     20
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1"
## [1] "Male first author team size 2018 geometric mean: 1.17346046000463"

## 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 = 55, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.0594630943593"
## [1] "Male last author team size 2018 geometric mean: 1.12246204830937"

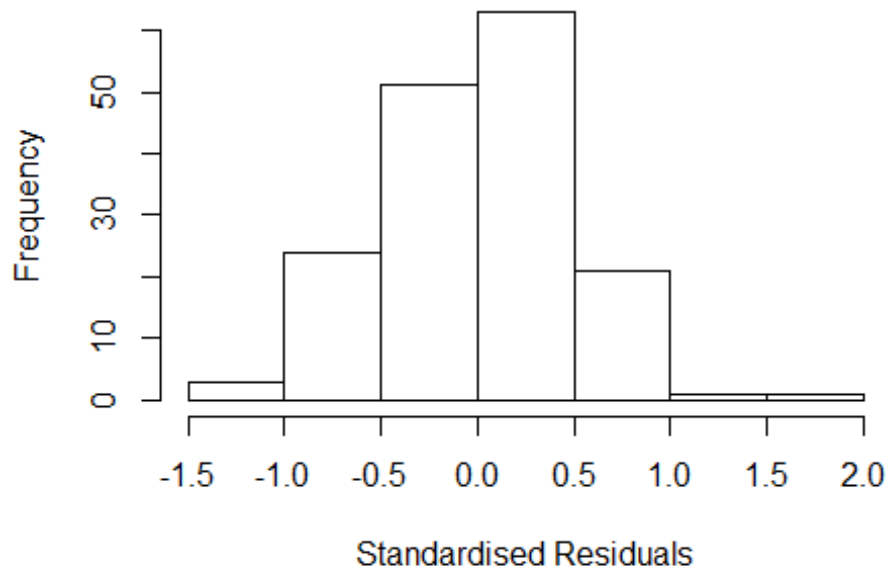
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 66, p-value = 0.6
## 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 8.296 1          2.880
## LastAuthorFemale 11.441 1          3.382
## UniqueAuthors    11.615 4          1.359
## Year              56.352 16         1.134
```

Residuals from first and last author and team size



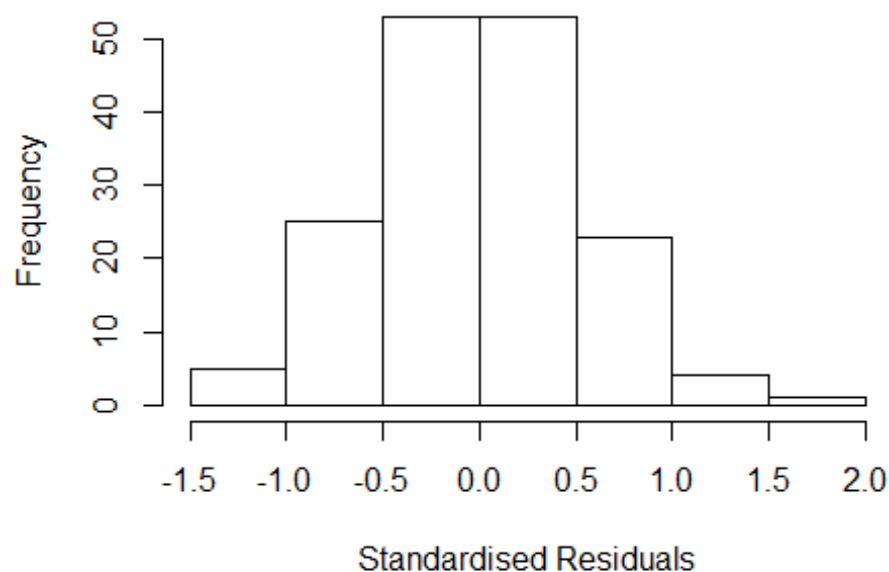
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3840 -0.3213 0.0298 0.3105 1.5509
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3652 0.2618 5.22 6.4e-07 ***
## FirstAuthorFemale1 0.1612 0.1126 1.43 0.15455
## LastAuthorFemale1 0.1587 0.1323 1.20 0.23241
## UniqueAuthors2 0.3558 0.1479 2.40 0.01748 *
## UniqueAuthors3 0.5832 0.1519 3.84 0.00019 ***
## UniqueAuthors4 0.8164 0.1728 4.72 5.5e-06 ***
## UniqueAuthors5 0.5251 0.1439 3.65 0.00037 ***
## Year1997 -0.8574 0.3179 -2.70 0.00785 **
## Year1998 0.0215 0.2661 0.08 0.93564
## Year1999 -0.2267 0.2987 -0.76 0.44913
```

```

## Year2000          -0.8519      0.3050    -2.79  0.00594 **
## Year2001          -0.4121      0.3444    -1.20  0.23355
## Year2002          -0.6272      0.2618    -2.40  0.01789 *
## Year2003          -0.3872      0.3148    -1.23  0.22078
## Year2004          -0.1764      0.2786    -0.63  0.52769
## Year2005          -0.5109      0.3100    -1.65  0.10160
## Year2006          -0.5358      0.2870    -1.87  0.06393 .
## Year2007          -0.5787      0.3114    -1.86  0.06525 .
## Year2008          -0.5483      0.2820    -1.94  0.05386 .
## Year2009          -0.5885      0.2714    -2.17  0.03183 *
## Year2010          -0.2508      0.3023    -0.83  0.40812
## Year2011          -0.5449      0.2761    -1.97  0.05043 .
## Year2012          -0.4980      0.2910    -1.71  0.08917 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.508
## Multiple R-squared:  0.326, Adjusted R-squared:  0.22
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 145 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.332  0.883  0.953  0.913  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          6.10e-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 9.913 1          3.148
## LastAuthorFemale 10.433 1          3.230
## Year              7.858 16          1.067

```

Residuals from first and last author



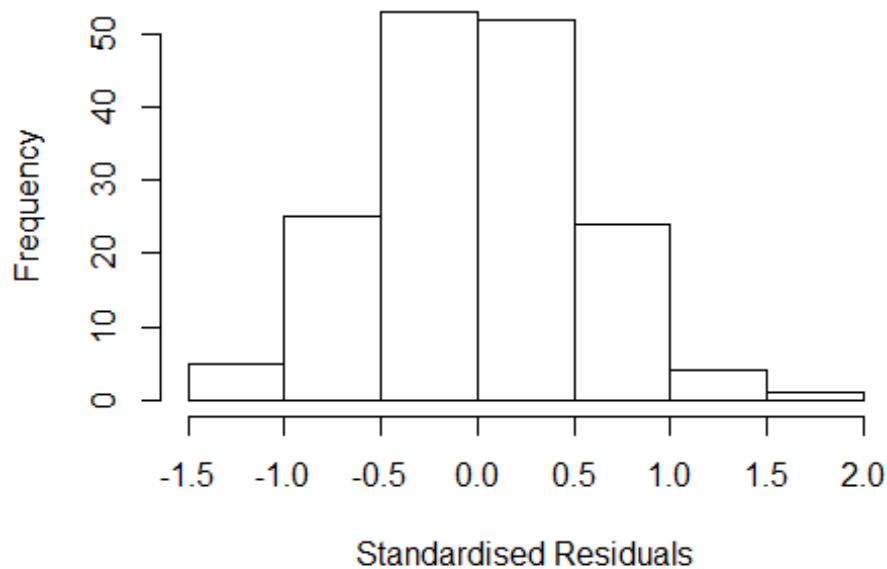
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2846 -0.3208 -0.0102 0.3300 1.7876
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4389 0.2533 5.68 7.1e-08 ***
## FirstAuthorFemale1 0.3143 0.1301 2.41 0.0170 *
## LastAuthorFemale1 0.0309 0.1358 0.23 0.8205
## Year1997 -0.6635 0.3650 -1.82 0.0712 .
## Year1998 -0.0776 0.2567 -0.30 0.7628
## Year1999 -0.1750 0.3051 -0.57 0.5673
## Year2000 -0.6988 0.3443 -2.03 0.0442 *
## Year2001 -0.3101 0.3333 -0.93 0.3537
## Year2002 -0.7009 0.2533 -2.77 0.0064 **
## Year2003 -0.1337 0.3303 -0.40 0.6862
## Year2004 -0.2584 0.2722 -0.95 0.3441
## Year2005 -0.2844 0.3337 -0.85 0.3954
```

```

## Year2006          -0.4499      0.2864   -1.57   0.1184
## Year2007          -0.5256      0.2964   -1.77   0.0783 .
## Year2008          -0.3839      0.2622   -1.46   0.1453
## Year2009          -0.5127      0.2776   -1.85   0.0668 .
## Year2010          -0.2733      0.3224   -0.85   0.3979
## Year2011          -0.4995      0.2727   -1.83   0.0691 .
## Year2012          -0.4731      0.2974   -1.59   0.1139
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.146, Adjusted R-squared:  0.0401
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 155 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.322  0.882  0.967  0.921  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      6.10e-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 4.867 1      2.206
## Year              4.867 16      1.051

```

Residuals from first author



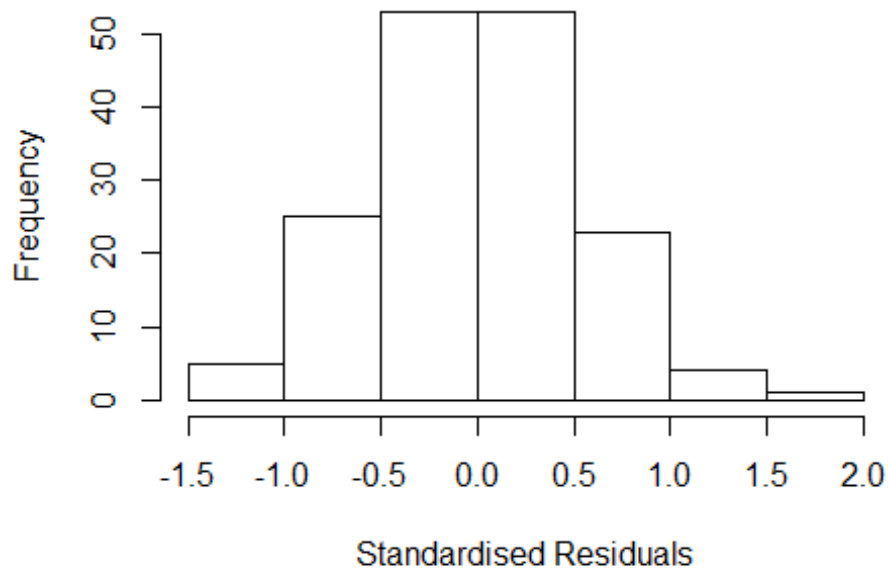
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.30441 -0.31774 -0.00835 0.33258 1.78891
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4523 0.2465 5.89 2.5e-08 ***
## FirstAuthorFemale1 0.3362 0.0911 3.69 0.00032 ***
## Year1997 -0.6814 0.3573 -1.91 0.05848 .
## Year1998 -0.0820 0.2582 -0.32 0.75124
## Year1999 -0.1848 0.3034 -0.61 0.54346
## Year2000 -0.7203 0.3352 -2.15 0.03330 *
## Year2001 -0.3255 0.3234 -1.01 0.31598
## Year2002 -0.7143 0.2465 -2.90 0.00434 **
## Year2003 -0.1553 0.3188 -0.49 0.62679
## Year2004 -0.2689 0.2683 -1.00 0.31788
## Year2005 -0.2847 0.3378 -0.84 0.40075
## Year2006 -0.4621 0.2814 -1.64 0.10266
```

```

## Year2007          -0.5322      0.2966   -1.79   0.07484 .
## Year2008          -0.3896      0.2623   -1.49   0.13956
## Year2009          -0.5238      0.2751   -1.90   0.05889 .
## Year2010          -0.2836      0.3286   -0.86   0.38961
## Year2011          -0.5142      0.2685   -1.91   0.05746 .
## Year2012          -0.4841      0.3006   -1.61   0.10943
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.561
## Multiple R-squared:  0.149, Adjusted R-squared:  0.0497
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 153 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.288  0.873  0.964  0.914  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      6.10e-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 5.763 1          2.401
## Year            5.763 16          1.056

```

Residuals from last author



```
## [1] "List of 0 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
## -1.29e+00 -3.35e-01 -5.35e-05 3.35e-01 1.69e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.44944 0.24083 6.02 1.4e-08 ***
## LastAuthorFemale1 0.25062 0.09729 2.58 0.0110 *
## Year1997 -0.57510 0.36349 -1.58 0.1158
## Year1998 0.00644 0.23727 0.03 0.9784
## Year1999 -0.15015 0.28720 -0.52 0.6019
## Year2000 -0.56983 0.35002 -1.63 0.1057
## Year2001 -0.27095 0.33787 -0.80 0.4239
## Year2002 -0.71144 0.24083 -2.95 0.0037 **
## Year2003 -0.03328 0.29429 -0.11 0.9101
## Year2004 -0.23755 0.25022 -0.95 0.3440
## Year2005 -0.28227 0.30324 -0.93 0.3535
## Year2006 -0.37547 0.28126 -1.33 0.1840
```

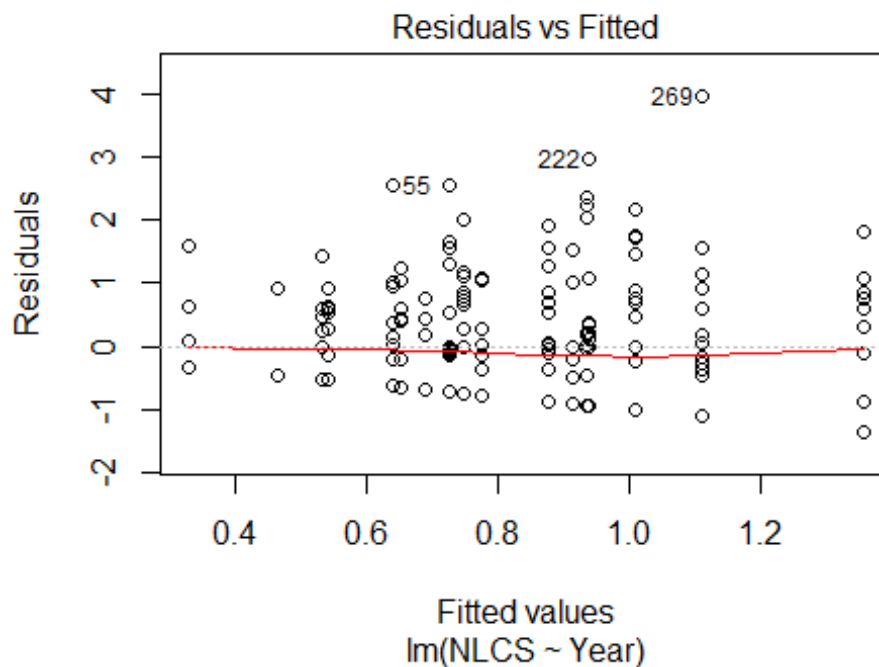


```

## Year2007          -0.54909      0.30370    -1.81    0.0727 .
## Year2008          -0.38936      0.24861    -1.57    0.1195
## Year2009          -0.45195      0.25913    -1.74    0.0832 .
## Year2010          -0.26916      0.30589    -0.88    0.3803
## Year2011          -0.41372      0.26006    -1.59    0.1138
## Year2012          -0.42837      0.28582    -1.50    0.1361
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.57
## Multiple R-squared:  0.113, Adjusted R-squared:  0.00936
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 148 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.358  0.865  0.955  0.911  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      6.10e-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: 164"
## [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
##    3    7    9   13    8   13   23   14   29   25   16   15   21   16   17
## 2011 2012
##   29   24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    9   12    7   12   21   13   24   22   14   12   20   16   15
## 2011 2012

```

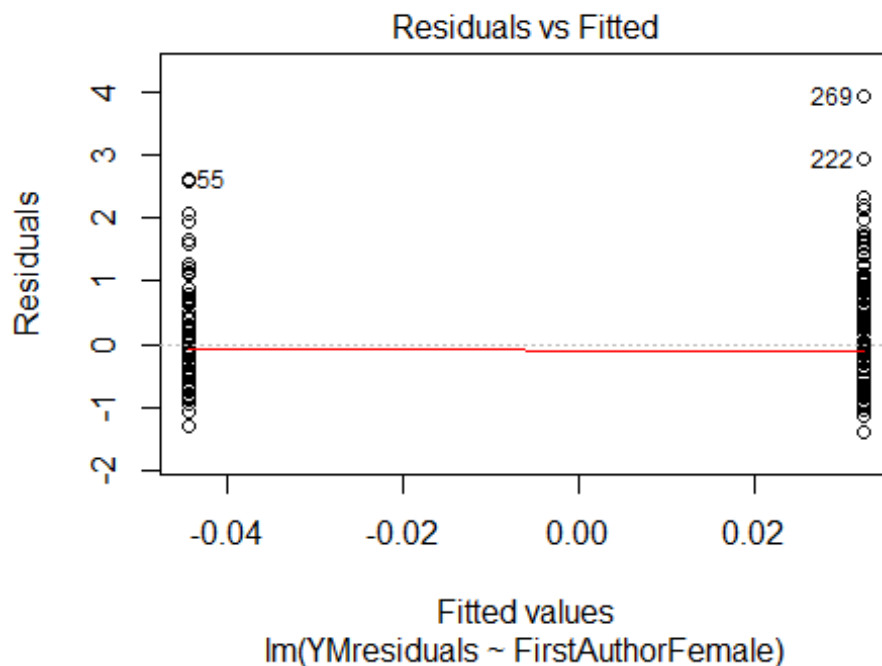
```
## 24 22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 5 9 12 7 12 21 13 24 22 14 12 20 16 15
## 2011 2012
## 24 22
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 1.2, df = 1, p-value = 0.3
## [1] "Female first author team size 2018 geometric mean: 1.14869835499704"
## [1] "Male first author team size 2018 geometric mean: 1"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

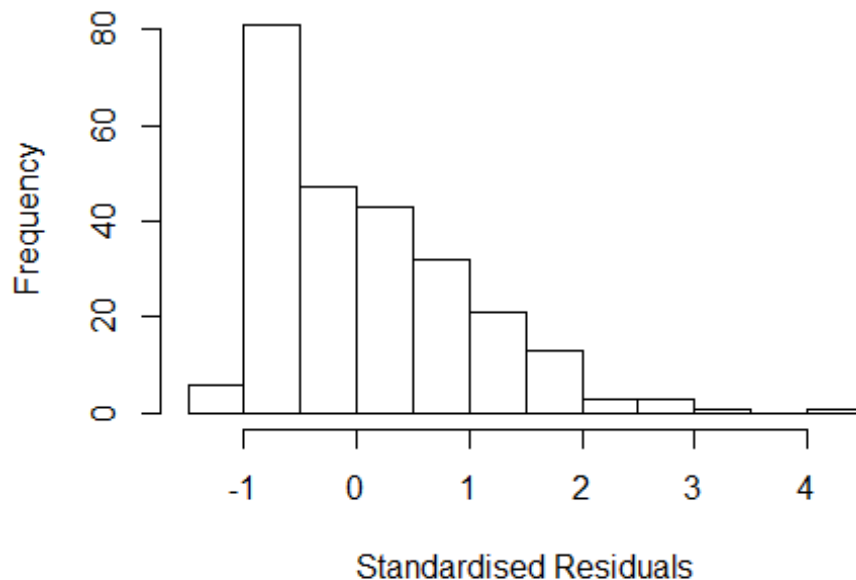
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 48, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 1.14869835499704"

## 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 = 0.2
## 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 -8.687e+14  1      NaN
## LastAuthorFemale  4.521e+01  1      6.724
## UniqueAuthors    -1.285e+16  3      NaN
## Year              -2.964e+15 16      NaN
```

Residuals from first and last author and team size



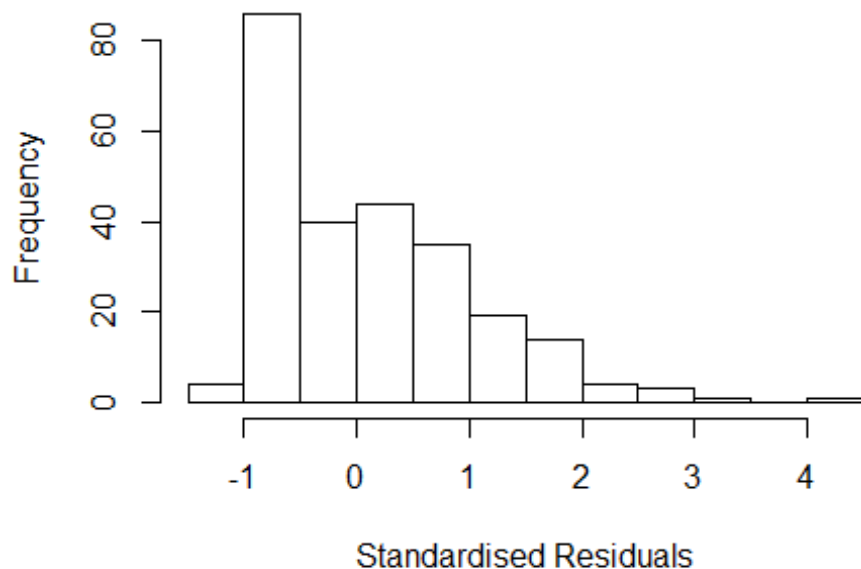
```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 208 70349839054 3.185 2009    1208     2    2.557
## 213 78650707059 3.306 2009    1208     1    2.678
## 222 79957891760 3.922 2010    1208     2    3.160
## 248 80053340024 3.290 2011    1208     2    2.809
## 269 84871015174 5.086 2012    1208     1    4.341
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.408 -0.547 -0.114  0.621  4.341
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4285     0.4239   1.01   0.313
## FirstAuthorFemale1 0.4505     0.5047   0.89   0.373
## LastAuthorFemale1 -0.4972     0.5123  -0.97   0.333
## UniqueAuthors2    0.5652     0.4484   1.26   0.209
## UniqueAuthors3    1.3117     0.2050   6.40 8.8e-10 ***
## UniqueAuthors4    0.4993     0.5671   0.88   0.380
```

```

## Year1997      0.2927      0.5200      0.56      0.574
## Year1998      0.2618      0.4932      0.53      0.596
## Year1999      0.2122      0.4779      0.44      0.657
## Year2000      0.4201      0.5667      0.74      0.459
## Year2001      0.1148      0.4551      0.25      0.801
## Year2002      0.1357      0.4549      0.30      0.766
## Year2003     -0.1279      0.4488     -0.29      0.776
## Year2004      0.1158      0.4657      0.25      0.804
## Year2005      0.4318      0.5256      0.82      0.412
## Year2006      0.9797      0.5809      1.69      0.093 .
## Year2007      0.0853      0.4641      0.18      0.854
## Year2008      0.3011      0.4648      0.65      0.518
## Year2009      0.1998      0.4861      0.41      0.681
## Year2010      0.3333      0.4810      0.69      0.489
## Year2011      0.0990      0.4624      0.21      0.831
## Year2012      0.3167      0.4901      0.65      0.519
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.762
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0306
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 238 is an outlier with |weight| = 0 ( < 0.0004);
## 7 weights are ~= 1. The remaining 243 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0473 0.8870 0.9480 0.8920 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      3.98e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 27.752 1      5.268
## LastAuthorFemale 29.129 1      5.397
## Year      2.556 16      1.030

```

Residuals from first and last author



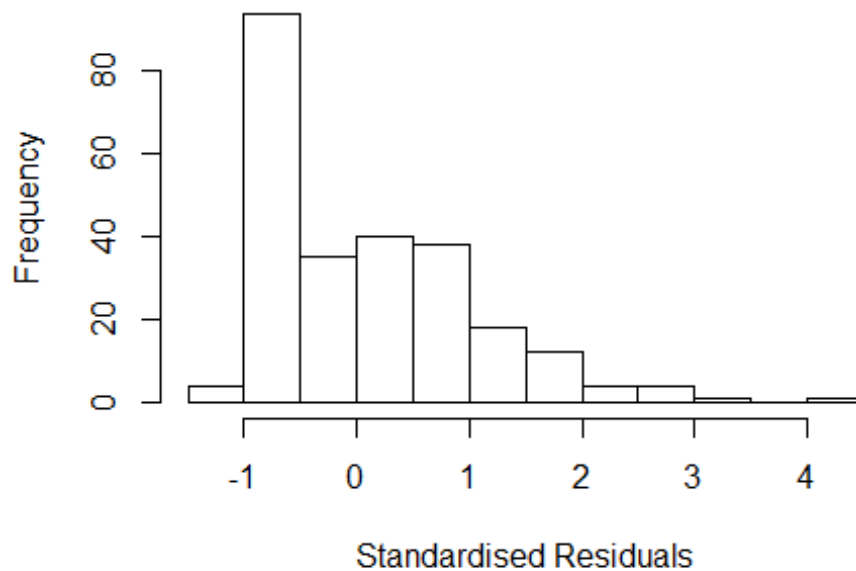
```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 208 70349839054 3.185 2009    1208     2    2.560
## 213 78650707059 3.306 2009    1208     1    2.681
## 222 79957891760 3.922 2010    1208     2    3.160
## 248 80053340024 3.290 2011    1208     2    2.822
## 269 84871015174 5.086 2012    1208     1    4.247
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.414 -0.572 -0.101  0.636  4.247
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4267    0.4258   1.00   0.317
## FirstAuthorFemale1 0.7704    0.5629   1.37   0.172
## LastAuthorFemale1 -0.8237    0.5715  -1.44   0.151
## Year1997         0.2986    0.5226   0.57   0.568
## Year1998         0.3369    0.4834   0.70   0.487
## Year1999         0.2154    0.4800   0.45   0.654
## Year2000         0.4207    0.5689   0.74   0.460
## Year2001         0.1178    0.4569   0.26   0.797
```

```

## Year2002          0.1356      0.4580      0.30      0.767
## Year2003         -0.1247      0.4503     -0.28      0.782
## Year2004          0.2254      0.4706      0.48      0.632
## Year2005          0.4263      0.5299      0.80      0.422
## Year2006          0.9876      0.5864      1.68      0.094 .
## Year2007          0.0881      0.4659      0.19      0.850
## Year2008          0.3162      0.4670      0.68      0.499
## Year2009          0.1985      0.4929      0.40      0.687
## Year2010          0.3354      0.4776      0.70      0.483
## Year2011          0.0942      0.4634      0.20      0.839
## Year2012          0.4127      0.5038      0.82      0.414
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.748
## Multiple R-squared:  0.0936, Adjusted R-squared:  0.0233
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## observation 238 is an outlier with |weight| = 0 ( < 0.0004);
## 9 weights are ~= 1. The remaining 241 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0349 0.8850 0.9410 0.8850 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          3.98e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.252 1          1.119
## Year              1.252 16          1.007

```

Residuals from first author



```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 208 70349839054 3.185 2009    1208     2    2.560
## 213 78650707059 3.306 2009    1208     1    2.681
## 222 79957891760 3.922 2010    1208     2    3.160
## 248 80053340024 3.290 2011    1208     2    2.822
## 269 84871015174 5.086 2012    1208     1    4.247
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3855 -0.5861 -0.0992  0.6410  4.1957
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.4309    0.4213     1.02   0.31
## FirstAuthorFemale1 -0.0151    0.1202    -0.13   0.90
## Year1997        0.2710    0.5165     0.52   0.60
## Year1998        0.3245    0.4796     0.68   0.50
## Year1999        0.1983    0.4749     0.42   0.68
## Year2000        0.4068    0.5637     0.72   0.47
## Year2001        0.1051    0.4529     0.23   0.82
## Year2002        0.0942    0.4503     0.21   0.83
```



```

## Year2003          -0.1452      0.4466    -0.33      0.75
## Year2004           0.2596      0.4647      0.56      0.58
## Year2005           0.4314      0.5182      0.83      0.41
## Year2006           0.9545      0.5772      1.65      0.10 .
## Year2007           0.0730      0.4612      0.16      0.87
## Year2008           0.3675      0.4658      0.79      0.43
## Year2009           0.1702      0.4912      0.35      0.73
## Year2010           0.3171      0.4790      0.66      0.51
## Year2011           0.0970      0.4603      0.21      0.83
## Year2012           0.4593      0.4813      0.95      0.34
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.784
## Multiple R-squared:  0.0704, Adjusted R-squared:  0.00253
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 238 is an outlier with |weight| = 0 ( < 0.0004);
## 12 weights are ~= 1. The remaining 238 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0645 0.8930 0.9420 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          3.98e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.359 1          1.166
## Year            1.359 16          1.010
##
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 208 70349839054 3.185 2009      1208      2      2.560
## 213 78650707059 3.306 2009      1208      1      2.681
## 222 79957891760 3.922 2010      1208      2      3.160
## 248 80053340024 3.290 2011      1208      2      2.822
## 269 84871015174 5.086 2012      1208      1      4.247

```

```
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.431 -0.581 -0.107  0.652  4.158
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.4322     0.4199   1.03   0.304
## LastAuthorFemale1 -0.0879     0.1240  -0.71   0.479
## Year1997          0.3130     0.5163   0.61   0.545
## Year1998          0.3460     0.4777   0.72   0.470
## Year1999          0.2287     0.4743   0.48   0.630
## Year2000          0.4486     0.5613   0.80   0.425
## Year2001          0.1221     0.4510   0.27   0.787
## Year2002          0.1297     0.4514   0.29   0.774
## Year2003         -0.1032     0.4469  -0.23   0.818
## Year2004          0.2876     0.4621   0.62   0.534
## Year2005          0.4611     0.5169   0.89   0.373
## Year2006          0.9986     0.5700   1.75   0.081 .
## Year2007          0.0998     0.4612   0.22   0.829
## Year2008          0.3795     0.4640   0.82   0.414
## Year2009          0.2363     0.4982   0.47   0.636
## Year2010          0.3656     0.4724   0.77   0.440
## Year2011          0.1316     0.4592   0.29   0.775
## Year2012          0.4957     0.4773   1.04   0.300
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.797
## Multiple R-squared:  0.0714, Adjusted R-squared:  0.00369
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 238 is an outlier with |weight| = 0 ( < 0.0004);
## 13 weights are ~= 1. The remaining 237 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.090  0.889  0.944  0.894  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      3.98e-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: 251"
## [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]"
##
## 2008 2009 2010 2011
##    3    2    1    1
##
## 2008 2009 2010 2011
##    1    2    0    0
##
## 2008 2009 2010 2011
##    1    2    0    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1"
## [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: 3"
## [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
##    2    1    1    2    3    1    2    1    4    5    1    9    4    6    1
## 2011 2012
##    6    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```

##      2      1      0      2      2      1      2      1      4      3      1      9      3      3      1
## 2011 2012
##      5      7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      0      2      2      1      2      1      4      3      1      9      3      3      1
## 2011 2012
##      4      7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.25992104989487"
## [1] "Male first 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 = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 1.25992104989487"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

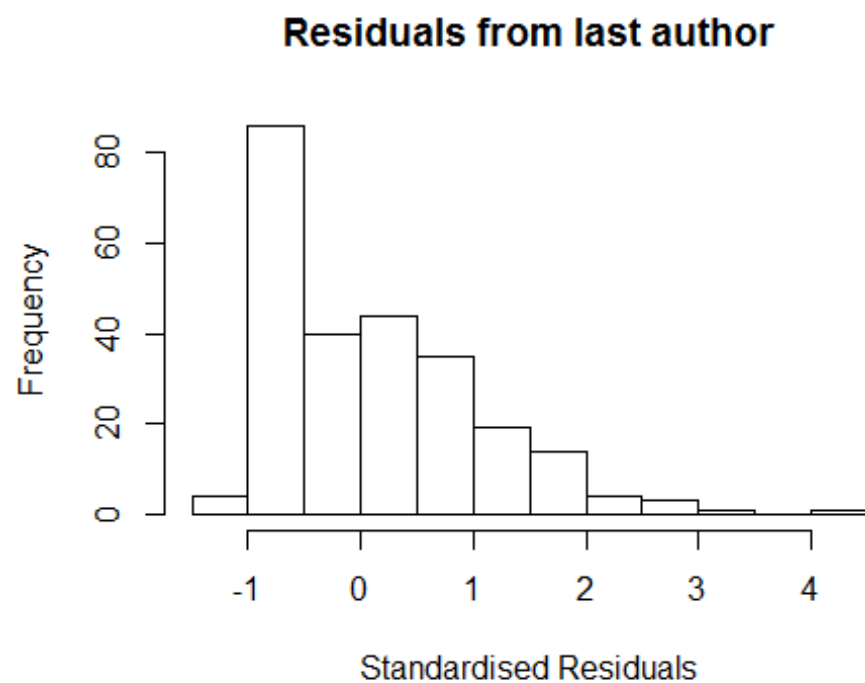
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## 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

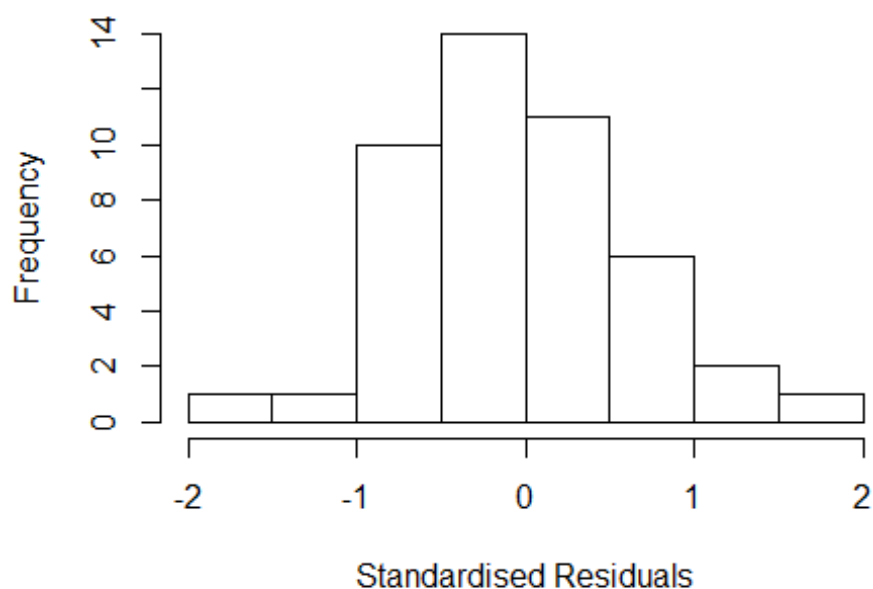
```



```
## [1] "Regression 2: First author gender, Last author gender, Year as factors"
```

	GVIF	Df	GVIF^(1/(2*Df))
## FirstAuthorFemale	-1.681e+15	1	NaN
## LastAuthorFemale	-1.090e+15	1	NaN
## Year	-2.717e+29	15	NaN

Residuals from first and last author



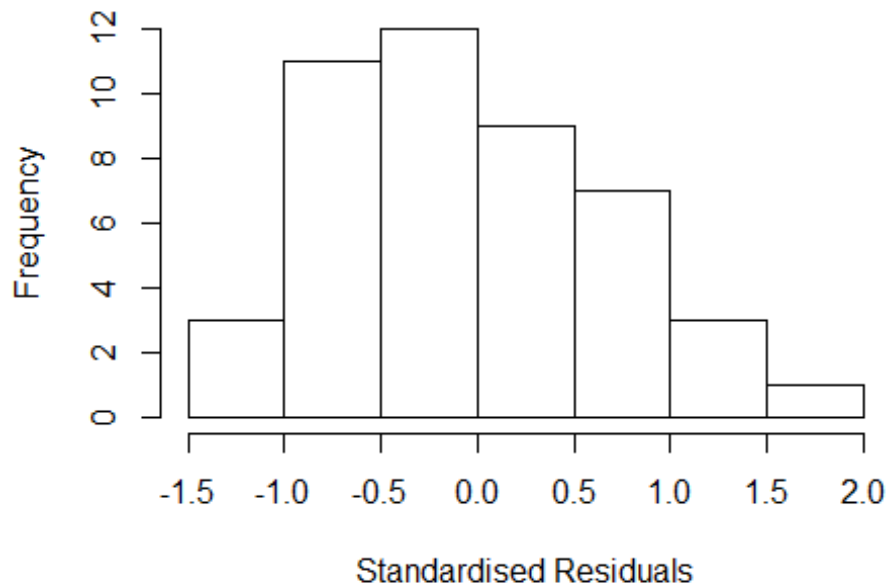
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.57e+00 -5.27e-01 3.61e-16 4.24e-01 1.76e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.0840 0.4340 4.80 4.8e-05 ***
## FirstAuthorFemale1 0.9266 0.8661 1.07 0.29382
## LastAuthorFemale1 -0.3939 0.8565 -0.46 0.64912
## Year1997 -2.0840 0.4340 -4.80 4.8e-05 ***
## Year1999 -1.3450 0.4500 -2.99 0.00578 **
## Year2000 -1.5213 0.6031 -2.52 0.01761 *
## Year2001 -1.6420 0.4340 -3.78 0.00075 ***
## Year2002 -1.6463 0.6694 -2.46 0.02036 *
## Year2003 -0.3046 0.9303 -0.33 0.74578
## Year2004 -1.6366 0.5700 -2.87 0.00770 **
## Year2005 -1.3130 0.5541 -2.37 0.02494 *
## Year2006 0.0423 0.5390 0.08 0.93799
```

```

## Year2007          -1.5476      0.5088   -3.04  0.00506 **
## Year2008          -1.2552      0.4876   -2.57  0.01563 *
## Year2009          -1.0411      0.9140   -1.14  0.26437
## Year2010          -0.6990      0.4340   -1.61  0.11845
## Year2011          -0.5098      1.3737   -0.37  0.71331
## Year2012          -1.3455      0.6051   -2.22  0.03441 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.76
## Multiple R-squared:  0.433, Adjusted R-squared:  0.0886
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.571  0.878  0.955  0.917  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      2.17e-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 1.085e+15  1      3.293e+07
## Year              1.085e+15 15      3.171e+00

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.42e+00 -5.14e-01 -8.17e-16  3.91e-01  1.71e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.0840     0.4256   4.90 3.4e-05 ***
## FirstAuthorFemale1  0.6056     0.2860   2.12 0.04287 *
## Year1997         -2.0840     0.4256  -4.90 3.4e-05 ***
## Year1999         -1.3450     0.4419  -3.04 0.00493 **
## Year2000         -1.5578     0.6055  -2.57 0.01547 *
## Year2001         -1.6420     0.4256  -3.86 0.00059 ***
## Year2002         -1.4858     0.6483  -2.29 0.02935 *
## Year2003          0.0164     0.5128   0.03 0.97478
## Year2004         -1.6756     0.5611  -2.99 0.00569 **
## Year2005         -1.3165     0.5377  -2.45 0.02063 *
## Year2006         -0.0306     0.5128  -0.06 0.95275
## Year2007         -1.5704     0.4975  -3.16 0.00371 **
```

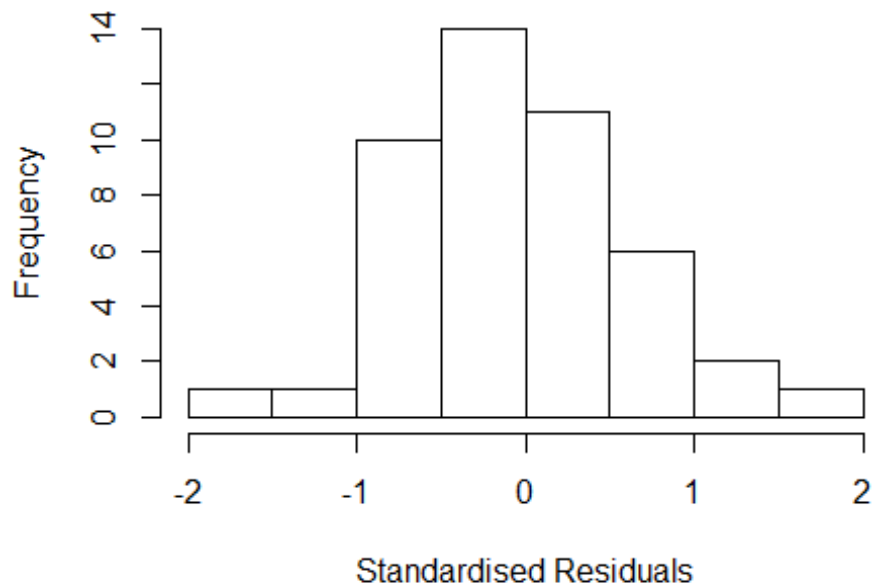


```

## Year2008          -1.2786      0.4706    -2.72  0.01100 *
## Year2009          -1.0674      0.8589    -1.24  0.22390
## Year2010          -0.6990      0.4256    -1.64  0.11133
## Year2011          -0.6654      0.9828    -0.68  0.50379
## Year2012          -1.2916      0.6008    -2.15  0.04003 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.836
## Multiple R-squared:  0.414, Adjusted R-squared:  0.0901
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.657  0.908  0.962  0.934  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      2.17e-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 8.714e+14 1      2.952e+07
## Year            8.714e+14 15      3.148e+00

```

Residuals from last author



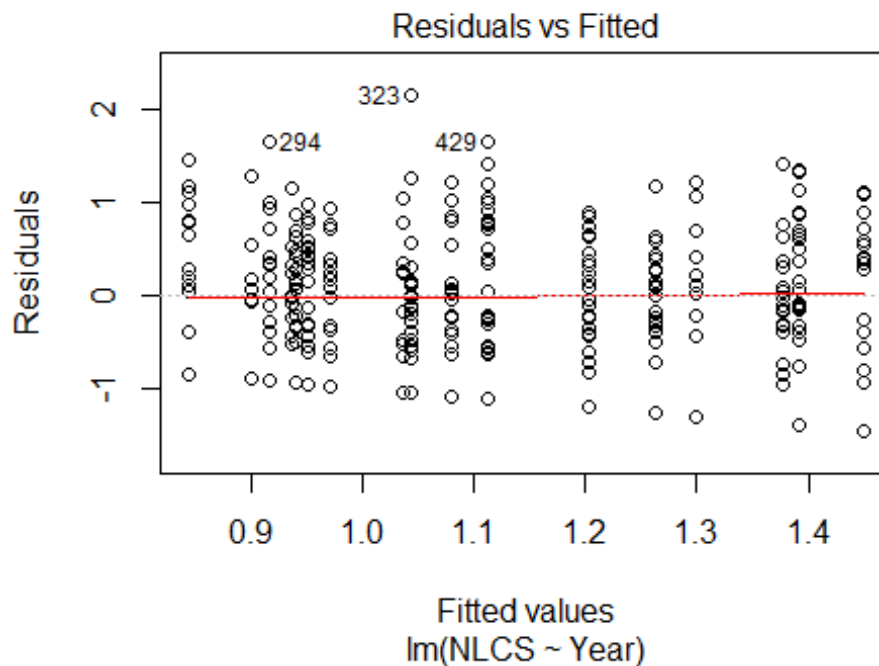
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.32e+00 -5.55e-01 -2.29e-16  4.44e-01  1.59e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.0840     0.4170   5.00 2.6e-05 ***
## LastAuthorFemale1  0.5016     0.2745   1.83 0.07795 .
## Year1997         -2.0840     0.4170  -5.00 2.6e-05 ***
## Year1999         -1.3450     0.4336  -3.10 0.00425 **
## Year2000         -1.5058     0.5662  -2.66 0.01261 *
## Year2001         -1.6420     0.4170  -3.94 0.00047 ***
## Year2002         -1.1830     0.8732  -1.35 0.18595
## Year2003          0.6220     0.4170   1.49 0.14658
## Year2004         -1.6293     0.5353  -3.04 0.00493 **
## Year2005         -1.3203     0.5275  -2.50 0.01820 *
## Year2006          0.0734     0.4992   0.15 0.88406
## Year2007         -1.5281     0.4861  -3.14 0.00383 **
```

```

## Year2008          -1.2446      0.4715    -2.64  0.01322 *
## Year2009          -1.0092      0.8096    -1.25  0.22256
## Year2010          -0.6990      0.4170    -1.68  0.10442
## Year2011          -0.7622      0.8306    -0.92  0.36637
## Year2012          -1.1788      0.6283    -1.88  0.07074 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.954
## Multiple R-squared:  0.375, Adjusted R-squared:  0.0295
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 35 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.762  0.916  0.959  0.939  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      2.17e-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: 46"
## [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
##   25   14   21   19   27   18   29   13   23   11   23   33   22   23   25
## 2011 2012
##   36   38
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   21   11   20   16   24   15   18   12   19   10   18   28   20   22   21
## 2011 2012
##   32   31

```

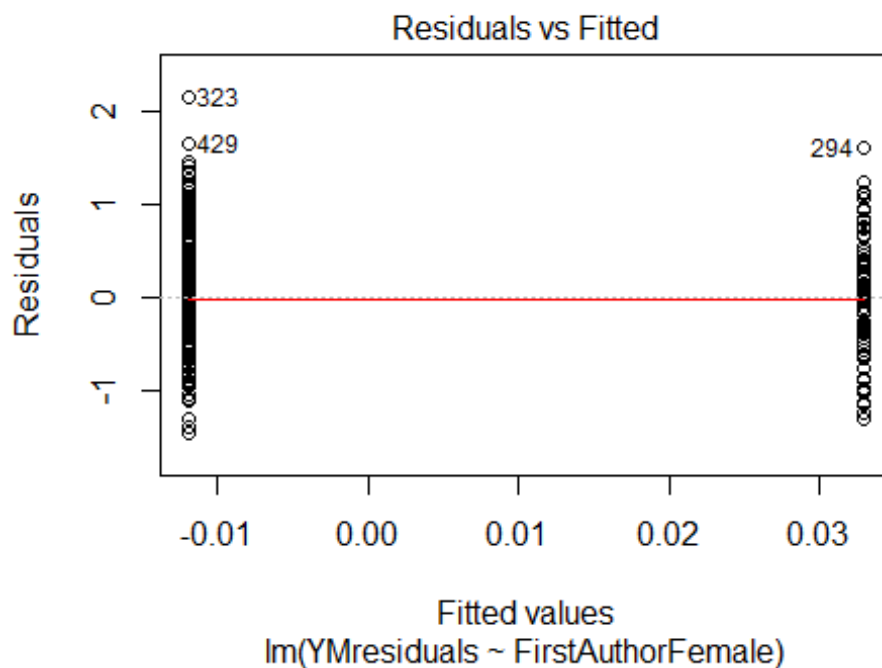
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 21 11 20 16 24 15 18 12 18 10 18 27 20 22 21
## 2011 2012
## 31 29
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS 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 = 0.28, df = 1, p-value = 0.6
## [1] "Female first author team size 2018 geometric mean: 1.63057279084246"
## [1] "Male first author team size 2018 geometric mean: 1.05476607648165"
## 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 = 0.004
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.47273335753469"
## [1] "Male last author team size 2018 geometric mean: 1.14280948374437"

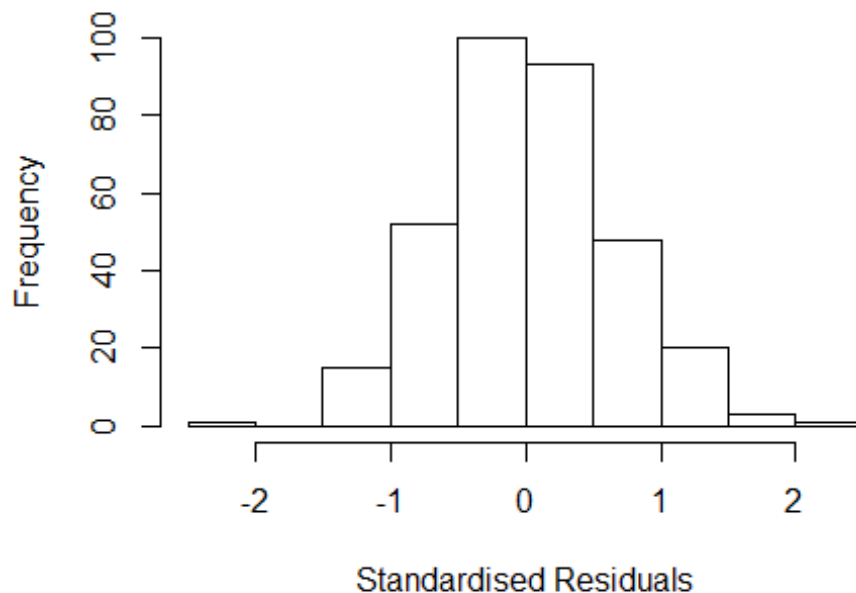
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170, p-value = 0.3
## 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	25.57	1	5.057
LastAuthorFemale	25.08	1	5.008
UniqueAuthors	22.85	3	1.684
Year	34.14	16	1.117

Residuals from first and last author and team size



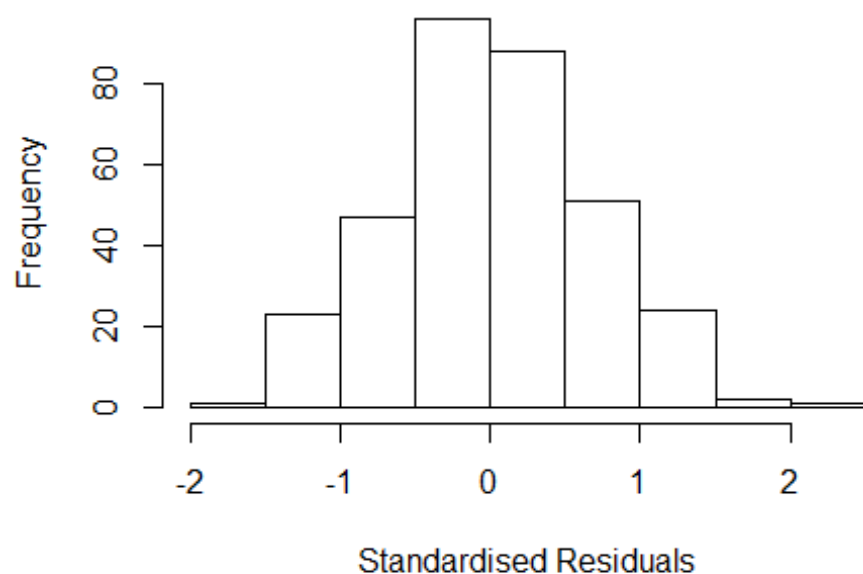
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -2.225984 -0.415816 -0.000072 0.420515 2.307654
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7499 0.2214 3.39 0.0008 ***
## FirstAuthorFemale1 -0.3428 0.3421 -1.00 0.3172
## LastAuthorFemale1 0.4362 0.3398 1.28 0.2002
## UniqueAuthors2 0.4414 0.1602 2.76 0.0062 **
## UniqueAuthors3 0.4559 0.1923 2.37 0.0184 *
## UniqueAuthors5 1.4349 0.1214 11.82 <2e-16 ***
## Year1997 0.2286 0.2857 0.80 0.4242
## Year1998 0.1819 0.2567 0.71 0.4791
## Year1999 0.1967 0.2623 0.75 0.4539
## Year2000 0.2060 0.2666 0.77 0.4404
```

```

## Year2001          0.1589      0.2453      0.65      0.5175
## Year2002          0.2860      0.2778      1.03      0.3042
## Year2003          0.5985      0.3570      1.68      0.0946 .
## Year2004          0.6078      0.2520      2.41      0.0164 *
## Year2005          0.0544      0.2895      0.19      0.8510
## Year2006          0.5316      0.3320      1.60      0.1103
## Year2007          0.3735      0.2428      1.54      0.1249
## Year2008          0.0662      0.2697      0.25      0.8062
## Year2009          0.1275      0.2459      0.52      0.6046
## Year2010          0.3559      0.2516      1.41      0.1582
## Year2011          0.5252      0.2510      2.09      0.0372 *
## Year2012          0.2101      0.3071      0.68      0.4944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.641
## Multiple R-squared:  0.139, Adjusted R-squared:  0.0808
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 300 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.168  0.858  0.950  0.903  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      3.00e-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 35.45 1      5.954
## LastAuthorFemale  34.58 1      5.881
## Year              1.69 16      1.017

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.67780 -0.44880 -0.00369 0.42270 2.25511
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7889 0.2166 3.64 0.00032 ***
## FirstAuthorFemale1 -0.2969 0.4840 -0.61 0.54009
## LastAuthorFemale1 0.3693 0.4828 0.77 0.44483
## Year1997 0.2330 0.2952 0.79 0.43051
## Year1998 0.1474 0.2517 0.59 0.55864
## Year1999 0.1628 0.2580 0.63 0.52842
## Year2000 0.1682 0.2611 0.64 0.51982
## Year2001 0.1227 0.2413 0.51 0.61136
## Year2002 0.2544 0.2737 0.93 0.35334
## Year2003 0.5196 0.3432 1.51 0.13106
## Year2004 0.6242 0.2559 2.44 0.01526 *
## Year2005 0.0773 0.2819 0.27 0.78405
```

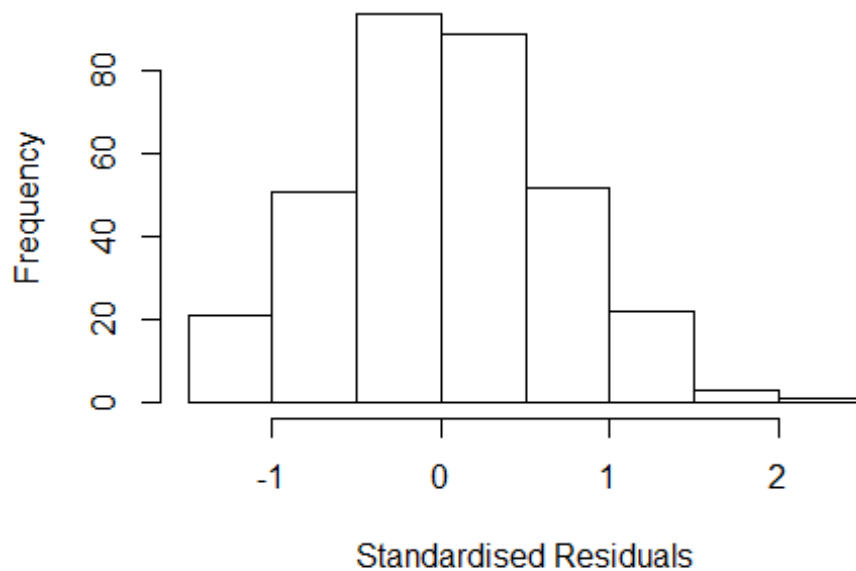


```

## Year2006          0.6662      0.3208      2.08  0.03865 *
## Year2007          0.3868      0.2438      1.59  0.11363
## Year2008          0.0644      0.2712      0.24  0.81246
## Year2009          0.1410      0.2399      0.59  0.55717
## Year2010          0.4384      0.2508      1.75  0.08148 .
## Year2011          0.5778      0.2557      2.26  0.02456 *
## Year2012          0.2259      0.3104      0.73  0.46730
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.659
## Multiple R-squared:  0.091, Adjusted R-squared:  0.0389
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 309 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.217  0.858  0.951  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      3.00e-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 1.252 1      1.119
## Year              1.252 16      1.007

```

Residuals from first author



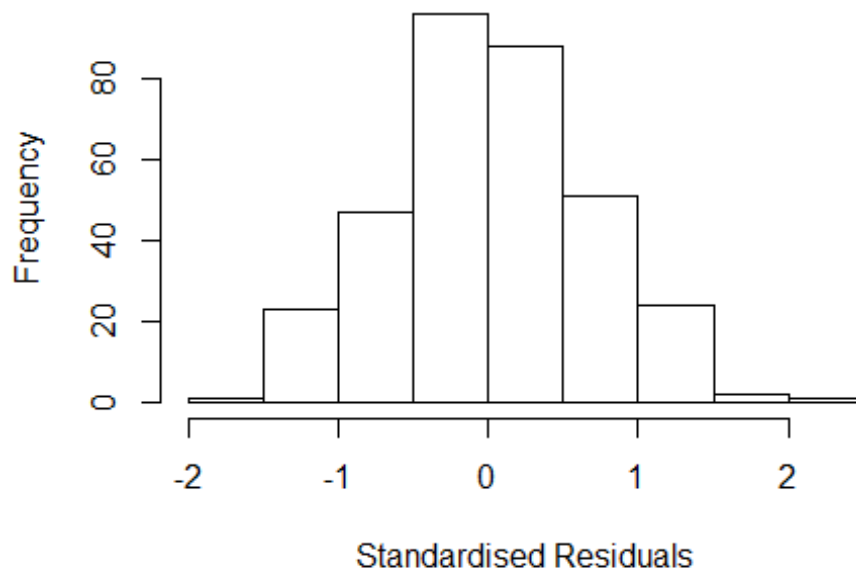
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.44255 -0.44962 0.00576 0.44627 2.25082
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7908 0.2178 3.63 0.00033 ***
## FirstAuthorFemale1 0.0588 0.0910 0.65 0.51858
## Year1997 0.2339 0.2960 0.79 0.43005
## Year1998 0.1489 0.2528 0.59 0.55619
## Year1999 0.1641 0.2590 0.63 0.52675
## Year2000 0.1678 0.2621 0.64 0.52253
## Year2001 0.1220 0.2424 0.50 0.61508
## Year2002 0.2566 0.2748 0.93 0.35111
## Year2003 0.5162 0.3309 1.56 0.11978
## Year2004 0.6018 0.2526 2.38 0.01779 *
## Year2005 0.0787 0.2836 0.28 0.78153
## Year2006 0.6517 0.3261 2.00 0.04651 *
```

```

## Year2007          0.3887      0.2448      1.59  0.11329
## Year2008          0.0815      0.2771      0.29  0.76884
## Year2009          0.1434      0.2411      0.59  0.55259
## Year2010          0.4609      0.2515      1.83  0.06777 .
## Year2011          0.5806      0.2570      2.26  0.02455 *
## Year2012          0.2179      0.3113      0.70  0.48432
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.0858, Adjusted R-squared:  0.0365
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 309 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.218  0.852  0.951  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      3.00e-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 1.24 1      1.113
## Year              1.24 16      1.007

```

Residuals from last author



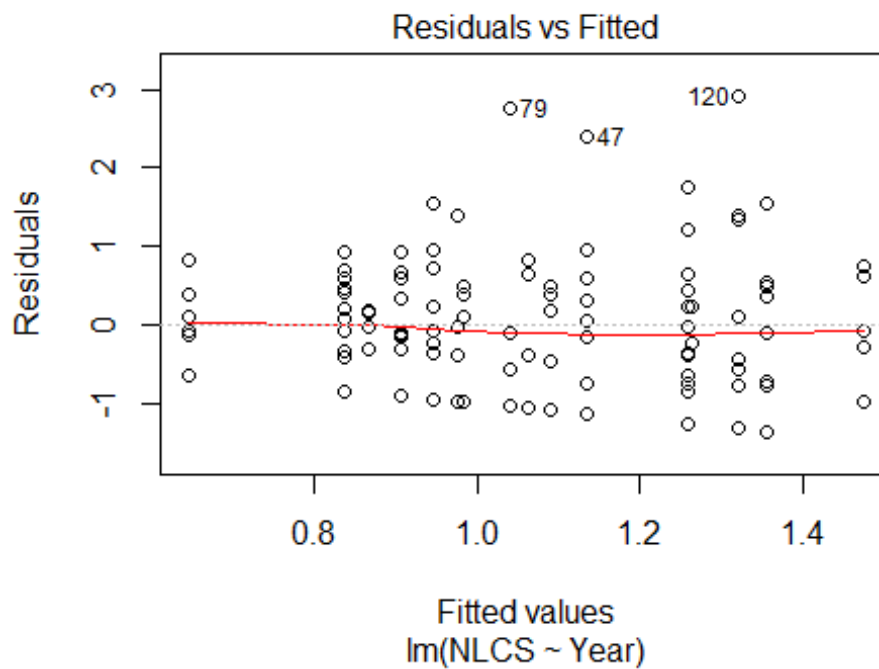
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.429703 -0.446896 -0.000919 0.424297 2.259018
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7870 0.2158 3.65 0.00031 ***
## LastAuthorFemale1 0.0840 0.0930 0.90 0.36733
## Year1997 0.2325 0.2948 0.79 0.43086
## Year1998 0.1463 0.2510 0.58 0.56039
## Year1999 0.1619 0.2573 0.63 0.52967
## Year2000 0.1689 0.2605 0.65 0.51728
## Year2001 0.1235 0.2405 0.51 0.60804
## Year2002 0.2528 0.2731 0.93 0.35530
## Year2003 0.5114 0.3301 1.55 0.12234
## Year2004 0.6042 0.2507 2.41 0.01652 *
## Year2005 0.0763 0.2807 0.27 0.78595
## Year2006 0.6427 0.3229 1.99 0.04744 *
```

```

## Year2007          0.3855      0.2432      1.58  0.11403
## Year2008          0.0759      0.2731      0.28  0.78134
## Year2009          0.1390      0.2391      0.58  0.56146
## Year2010          0.4526      0.2498      1.81  0.07103 .
## Year2011          0.5756      0.2549      2.26  0.02460 *
## Year2012          0.2188      0.3087      0.71  0.47899
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.658
## Multiple R-squared:  0.0869, Adjusted R-squared:  0.0377
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 310 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.215  0.857  0.951  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      3.00e-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: 333"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4   10    2    5    6    5    6    7   10   12    9    9   14   11   13
## 2012
##    13
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4   10    2    4    5    4    4    6   10   12    8    5   13    9   10
## 2012

```

```
## 12
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 4 10 2 4 5 4 4 6 10 12 8 5 13 9 10
## 2012
## 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 = 24, df = 15, p-value = 0.06
```



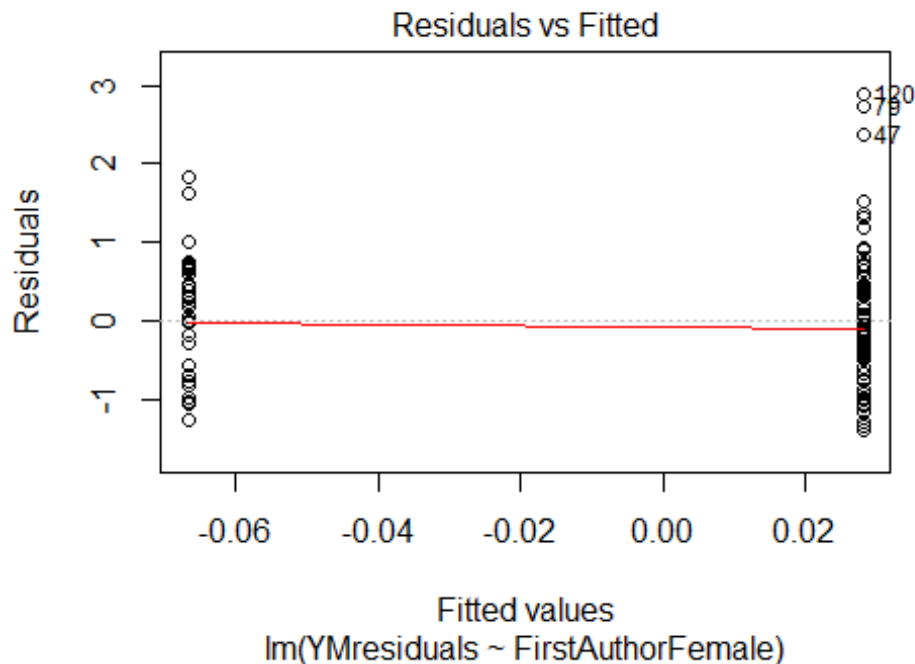
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.79, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 1.73205080756888"
## [1] "Male first author team size 2018 geometric mean: 1.11253147609649"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

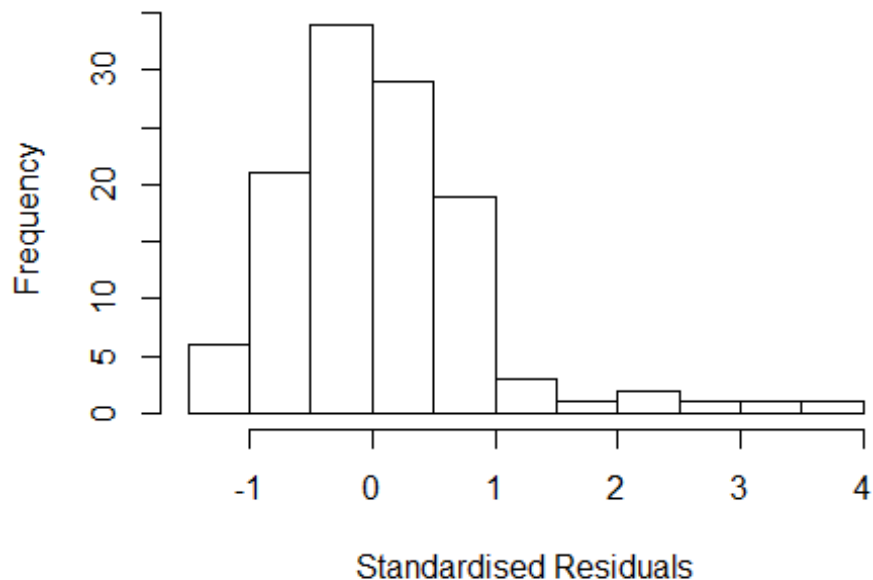
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.44224957030741"
## [1] "Male last author team size 2018 geometric mean: 1.19422009496884"

## 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 = 24, 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 1.147e+01  1      3.387
## LastAuthorFemale  9.923e+00  1      3.150
## UniqueAuthors    1.200e+17  3     702.317
## Year              1.587e+17 15      3.744
```

Residuals from first and last author and team size



```
## [1] "List of 3 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 47  33845544068 3.537 2005     1212      1      2.572
## 79  61149729209 3.805 2008     1212      1      3.434
## 120 84859029690 4.231 2011     1212      1      3.608
##
## 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
## -1.2850 -0.4193 -0.0133  0.4835  3.6080
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8694    0.1117   7.78 7.8e-12 ***
## FirstAuthorFemale1 -0.3629    0.2132  -1.70  0.092 .
## LastAuthorFemale1  0.3554    0.2072   1.72  0.090 .
## UniqueAuthors2    0.3595    0.3278   1.10  0.276
## UniqueAuthors4    1.4770    0.2449   6.03 3.0e-08 ***
## UniqueAuthors5    2.0870    0.3386   6.16 1.6e-08 ***
## Year1998          0.0413    0.2246   0.18  0.855
## Year1999          0.3973    0.2014   1.97  0.051 .
```

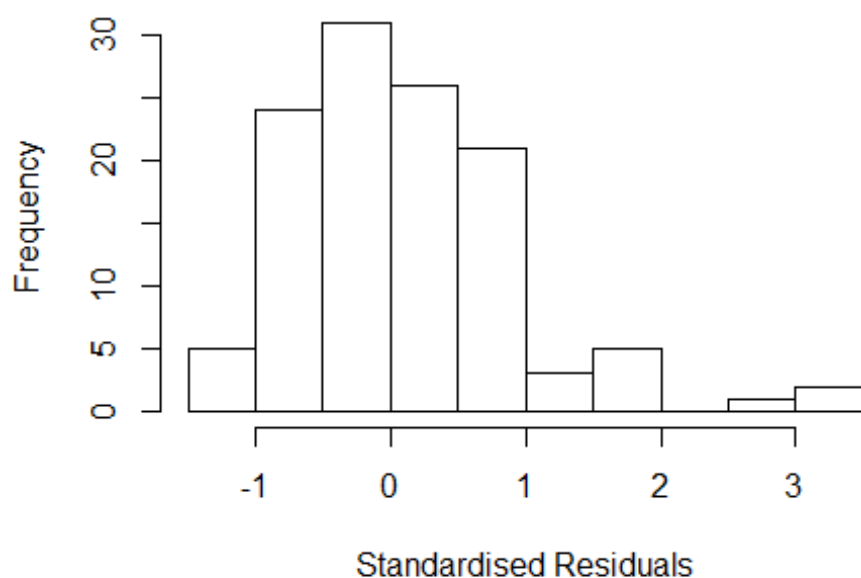


```

## Year2000          0.1467      0.3212      0.46      0.649
## Year2001          0.6190      0.3239      1.91      0.059 .
## Year2002          0.2149      0.4635      0.46      0.644
## Year2003          0.0337      0.4958      0.07      0.946
## Year2004          0.2501      0.2920      0.86      0.394
## Year2005          0.0955      0.3760      0.25      0.800
## Year2006          0.0378      0.2399      0.16      0.875
## Year2007          0.4155      0.3448      1.21      0.231
## Year2008         -0.4980      0.2657     -1.87      0.064 .
## Year2009         -0.0526      0.1983     -0.27      0.791
## Year2010         -0.2279      0.2247     -1.01      0.313
## Year2011         -0.2464      0.3511     -0.70      0.484
## Year2012          0.1306      0.2761      0.47      0.637
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.802
## Multiple R-squared:  0.185, Adjusted R-squared:  0.0171
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0061 0.8880 0.9560 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          8.47e-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 3.400 1          1.844
## LastAuthorFemale  3.690 1          1.921
## Year              1.456 15          1.013

```

Residuals from first and last author



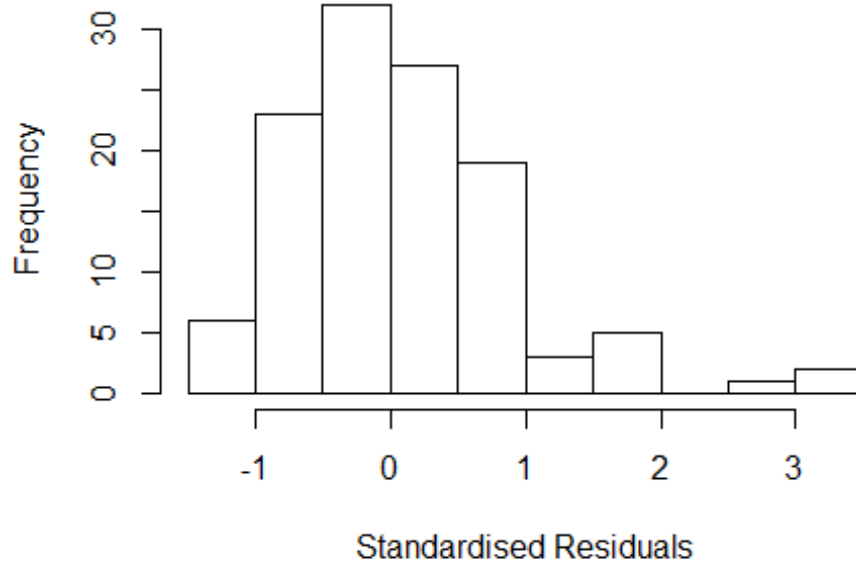
```
## [1] "List of 3 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 47  33845544068 3.537 2005    1212     1    2.557
## 79  61149729209 3.805 2008    1212     1    3.407
## 120 84859029690 4.231 2011    1212     1    3.314
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3468 -0.4191 -0.0303  0.5240  3.4067
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8706    0.1119   7.78 6.8e-12 ***
## FirstAuthorFemale1 -0.2661    0.2445  -1.09  0.279
## LastAuthorFemale1  0.2538    0.2627   0.97  0.336
## Year1998         0.0412    0.2242   0.18  0.855
## Year1999         0.3986    0.2000   1.99  0.049 *
## Year2000         0.1445    0.3205   0.45  0.653
## Year2001         0.6191    0.3224   1.92  0.058 .
## Year2002         0.2134    0.4586   0.47  0.643
## Year2003         0.0409    0.4935   0.08  0.934
## Year2004         0.2937    0.2915   1.01  0.316
```

```

## Year2005          0.1093      0.3560      0.31      0.760
## Year2006          0.0418      0.2405      0.17      0.862
## Year2007          0.4762      0.3537      1.35      0.181
## Year2008         -0.4722      0.2882     -1.64      0.104
## Year2009         -0.0248      0.2026     -0.12      0.903
## Year2010         -0.2269      0.2242     -1.01      0.314
## Year2011          0.0464      0.4783      0.10      0.923
## Year2012          0.3357      0.2716      1.24      0.219
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.832
## Multiple R-squared:  0.088, Adjusted R-squared:  -0.067
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0556 0.8940 0.9560 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      8.47e-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 1.306 1      1.143
## Year              1.306 15      1.009

```

Residuals from first author



```
## [1] "List of 3 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 47  33845544068 3.537 2005     1212      1      2.557
## 79  61149729209 3.805 2008     1212      1      3.407
## 120 84859029690 4.231 2011     1212      1      3.314
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3534 -0.4754 -0.0271  0.5004  3.3912
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8761    0.1130   7.75 7.4e-12 ***
## FirstAuthorFemale1 -0.0345    0.1565  -0.22  0.826
## Year1998        0.0413    0.2259   0.18  0.855
## Year1999        0.4041    0.1941   2.08  0.040 *
## Year2000        0.1446    0.3213   0.45  0.654
## Year2001        0.6230    0.3244   1.92  0.058 .
## Year2002        0.2134    0.4615   0.46  0.645
## Year2003        0.0447    0.4936   0.09  0.928
## Year2004        0.2491    0.2670   0.93  0.353
## Year2005        0.1151    0.3707   0.31  0.757
```

```

## Year2006          0.0463      0.2411      0.19      0.848
## Year2007          0.4773      0.3535      1.35      0.180
## Year2008         -0.4623      0.2937     -1.57      0.119
## Year2009         -0.0255      0.2038     -0.12      0.901
## Year2010         -0.2246      0.2249     -1.00      0.320
## Year2011          0.0570      0.4762      0.12      0.905
## Year2012          0.3385      0.2861      1.18      0.240
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.841
## Multiple R-squared:  0.0841, Adjusted R-squared:  -0.061
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0669 0.8940 0.9560 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      8.47e-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 1.438 1      1.199
## Year      1.438 15      1.012

## [1] "List of 3 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 47 33845544068 3.537 2005      1212      1      2.557
## 79 61149729209 3.805 2008      1212      1      3.407
## 120 84859029690 4.231 2011      1212      1      3.314
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"

```

```

## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3422 -0.4663 -0.0288  0.5206  3.4160
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.86669    0.11041    7.85 4.5e-12 ***
## LastAuthorFemale1 0.00316    0.16545    0.02  0.985
## Year1998        0.04112    0.22304    0.18  0.854
## Year1999        0.39473    0.20383    1.94  0.056 .
## Year2000        0.14432    0.31991    0.45  0.653
## Year2001        0.61637    0.32076    1.92  0.057 .
## Year2002        0.21334    0.45635    0.47  0.641
## Year2003        0.03837    0.49272    0.08  0.938
## Year2004        0.24928    0.27503    0.91  0.367
## Year2005        0.10564    0.35706    0.30  0.768
## Year2006        0.03880    0.23950    0.16  0.872
## Year2007        0.47547    0.35390    1.34  0.182
## Year2008       -0.47774    0.28425   -1.68  0.096 .
## Year2009       -0.02433    0.20187   -0.12  0.904
## Year2010       -0.22852    0.22353   -1.02  0.309
## Year2011        0.04014    0.47725    0.08  0.933
## Year2012        0.33546    0.28348    1.18  0.239
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.827
## Multiple R-squared:  0.086, Adjusted R-squared:  -0.0588
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
##  11 weights are ~= 1. The remaining 107 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0494 0.8930 0.9550 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      8.47e-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)

```

```

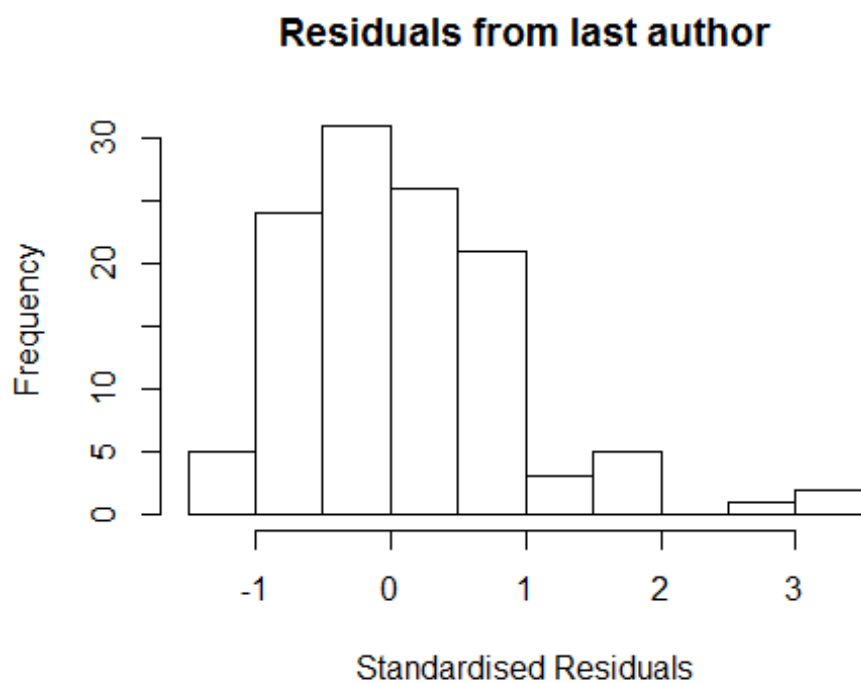
## [1] "Sample size for the above analysis: 118"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1213"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    3    4    2    4    7    5    4    5    7   16   24   14   48   27
##
## 1997 1998 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    3    2    1    4    6    5    4    5    7   14   23   13   44   25
##
## 1997 1998 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    3    2    1    4    4    5    4    3    7   13   20   12   44   24
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.3152568189817"
## [1] "Male first author team size 2018 geometric mean: 1.13431252219546"

## 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 = 120, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.13011578342933"
## [1] "Male last author team size 2018 geometric mean: 1.40770919098948"

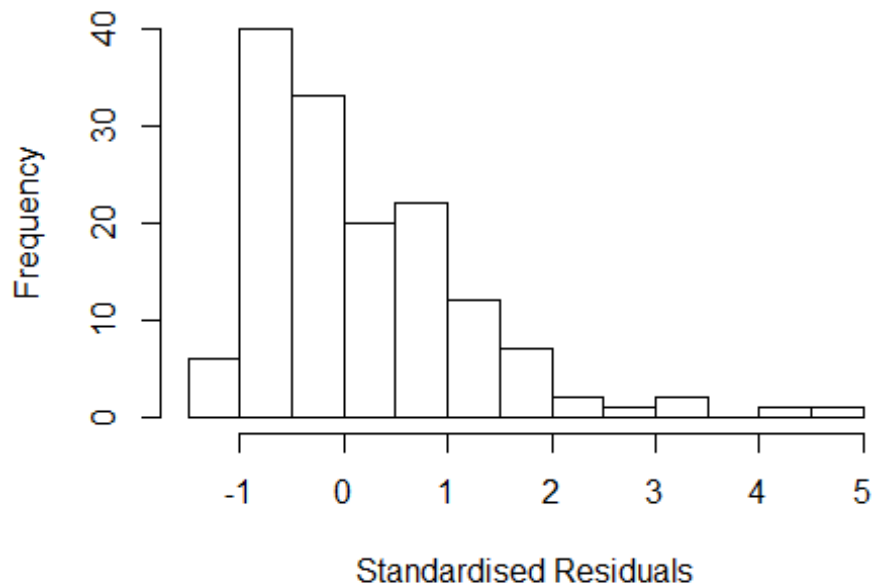
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 94, 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 9.297e+14  1      3.049e+07
## LastAuthorFemale  1.103e+02  1      1.050e+01
## UniqueAuthors    5.216e+14  1      2.284e+07
## Year              7.500e+15 14      3.689e+00
```


Residuals from first and last author and team size



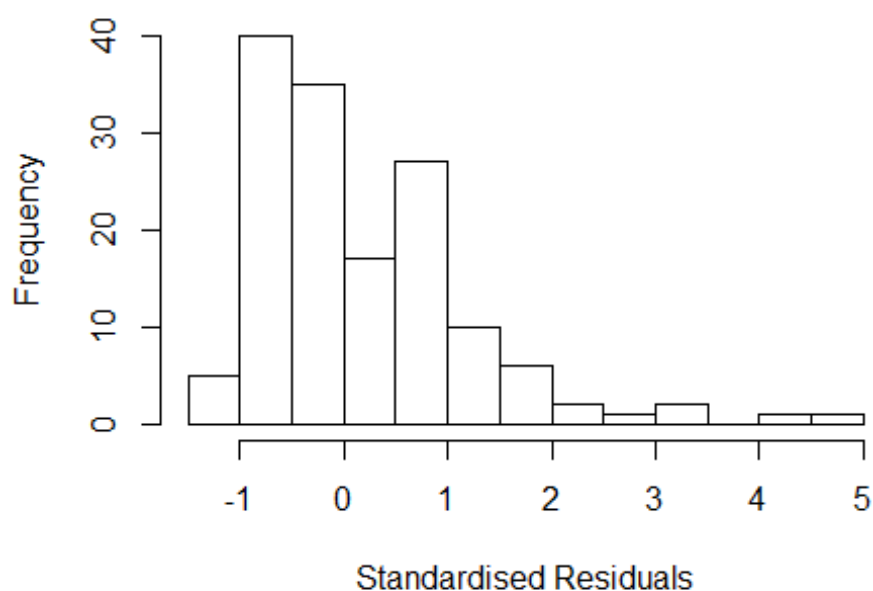
```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 86  77957886942 4.361 2009    1213     1    4.062
## 95  79957891760 3.922 2010    1208     2    3.043
## 130 80053340024 3.290 2011    1208     2    2.596
## 158 84870719154 4.195 2012    1213     1    3.446
## 164 84864343442 5.324 2012    1213     1    4.575
##
## 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
## -1.412 -0.620 -0.100  0.668  4.575
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.3707     0.3955   3.47 0.00072 ***
## FirstAuthorFemale1  0.3871     0.5413   0.72 0.47588
## LastAuthorFemale1 -0.3130     0.5062  -0.62 0.53744
## UniqueAuthors2    0.4303     0.4055   1.06 0.29058
## Year1998          -0.8498     0.7069  -1.20 0.23151
## Year1999          -0.2744     0.2218  -1.24 0.21844
```

```

## Year2001          -1.4448      0.4087    -3.54  0.00057 ***
## Year2002           0.3460      0.6193     0.56  0.57741
## Year2003          -0.8010      0.3959    -2.02  0.04509 *
## Year2004           0.0413      0.6884     0.06  0.95220
## Year2005           0.3316      0.5314     0.62  0.53374
## Year2006           0.1407      0.3966     0.35  0.72342
## Year2007          -0.3054      0.5082    -0.60  0.54903
## Year2008           0.0100      0.5804     0.02  0.98622
## Year2009          -1.0721      0.4134    -2.59  0.01060 *
## Year2010          -0.4921      0.4448    -1.11  0.27065
## Year2011          -0.7508      0.3733    -2.01  0.04639 *
## Year2012          -0.6218      0.4371    -1.42  0.15733
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.784
## Multiple R-squared:  0.193, Adjusted R-squared:  0.0863
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 2 observations c(65,132) are outliers with |weight| = 0 ( < 0.00068);
## 8 weights are ~= 1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0142 0.8700 0.9350 0.8880 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          6.80e-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 1.009e+13 1      3.176e+06
## LastAuthorFemale  6.057e+01 1      7.782e+00
## Year              1.044e+14 14      3.167e+00

```

Residuals from first and last author



```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 86  77957886942 4.361 2009    1213     1    4.046
## 95  79957891760 3.922 2010    1208     2    3.022
## 130 80053340024 3.290 2011    1208     2    2.545
## 158 84870719154 4.195 2012    1213     1    3.406
## 164 84864343442 5.324 2012    1213     1    4.535
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.424 -0.685 -0.106  0.723  4.535
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.423958   0.391749   3.63  0.00040 ***
## FirstAuthorFemale1 -0.004390   0.402979  -0.01  0.99132
## LastAuthorFemale1  0.064042   0.391749  0.16  0.87040
## Year1998        -0.892490   0.699688  -1.28  0.20439
## Year1999        -0.300979   0.229551  -1.31  0.19211
## Year2001        -1.483610   0.402979  -3.68  0.00034 ***
## Year2002         0.322868   0.621025  0.52  0.60402
## Year2003        -0.850642   0.391195  -2.17  0.03148 *
```

```

## Year2004          0.000541    0.676748    0.00  0.99936
## Year2005          0.291031    0.524449    0.55  0.57990
## Year2006          0.092217    0.392491    0.23  0.81461
## Year2007         -0.344812    0.504062   -0.68  0.49515
## Year2008         -0.030075    0.568773   -0.05  0.95791
## Year2009         -1.109060    0.412311   -2.69  0.00809 **
## Year2010         -0.524235    0.443230   -1.18  0.23906
## Year2011         -0.738803    0.370839   -1.99  0.04844 *
## Year2012         -0.634661    0.443069   -1.43  0.15442
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.821
## Multiple R-squared:  0.18,    Adjusted R-squared:  0.0786
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 2 observations c(65,132) are outliers with |weight| = 0 ( < 0.00068);
## 8 weights are ~= 1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0466  0.8850  0.9360  0.8960  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          6.80e-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 1.209e+12  1          1.099e+06
## Year              1.209e+12 14          2.701e+00

## [1] "List of 5 outliers with residuals above 2.5"
##           ScopusId  NLCS Year OneField Fields residuals
## 86  77957886942  4.361 2009    1213      1    4.046
## 95  79957891760  3.922 2010    1208      2    3.022
## 130 80053340024  3.290 2011    1208      2    2.545
## 158 84870719154  4.195 2012    1213      1    3.406
## 164 84864343442  5.324 2012    1213      1    4.535
##

```

```

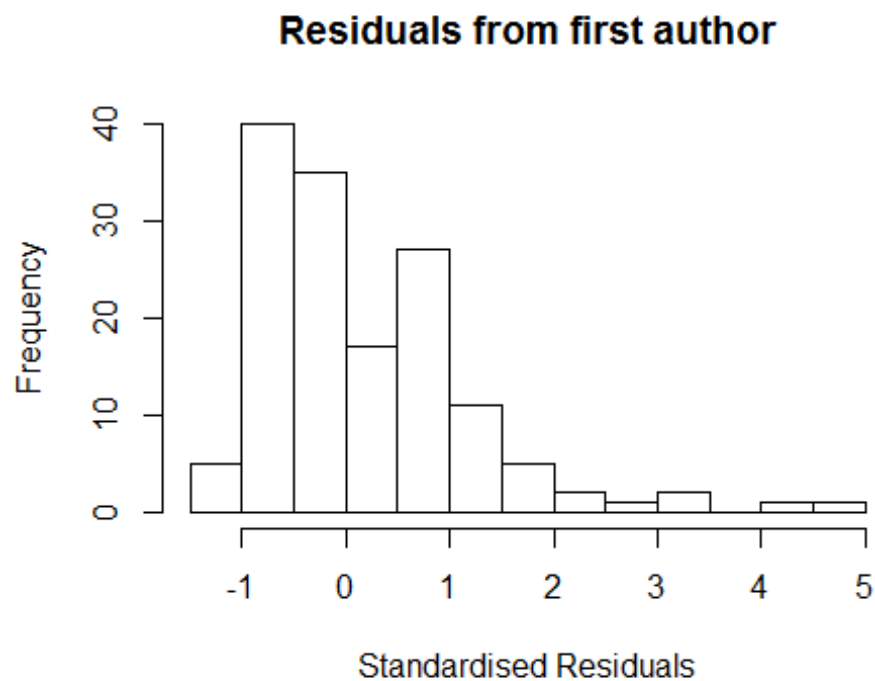
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.427 -0.694 -0.107  0.721  4.530
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.49e+00   1.69e-07  8.81e+06 < 2e-16 ***
## FirstAuthorFemale1 5.55e-02   1.66e-01  3.30e-01  0.7390
## Year1998        -9.55e-01   5.91e-01 -1.62e+00  0.1084
## Year1999        -3.33e-01   1.43e-01 -2.33e+00  0.0214 *
## Year2001        -1.54e+00   1.66e-01 -9.28e+00  4.6e-16 ***
## Year2002         2.65e-01   5.21e-01  5.10e-01  0.6116
## Year2003        -9.14e-01   6.57e-02 -1.39e+01 < 2e-16 ***
## Year2004        -6.05e-02   5.69e-01 -1.10e-01  0.9154
## Year2005         2.31e-01   3.69e-01  6.30e-01  0.5327
## Year2006         2.95e-02   8.39e-02  3.50e-01  0.7254
## Year2007        -4.05e-01   3.46e-01 -1.17e+00  0.2433
## Year2008        -9.13e-02   4.34e-01 -2.10e-01  0.8335
## Year2009        -1.17e+00   1.91e-01 -6.11e+00  1.0e-08 ***
## Year2010        -5.84e-01   2.48e-01 -2.36e+00  0.0199 *
## Year2011        -7.94e-01   1.92e-01 -4.14e+00  6.2e-05 ***
## Year2012        -6.94e-01   2.63e-01 -2.64e+00  0.0093 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.828
## Multiple R-squared:  0.178, Adjusted R-squared:  0.0844
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 2 observations c(65,132) are outliers with |weight| = 0 ( < 0.00068);
## 8 weights are ~ = 1. The remaining 137 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.0534 0.8870 0.9370 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      6.80e-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"

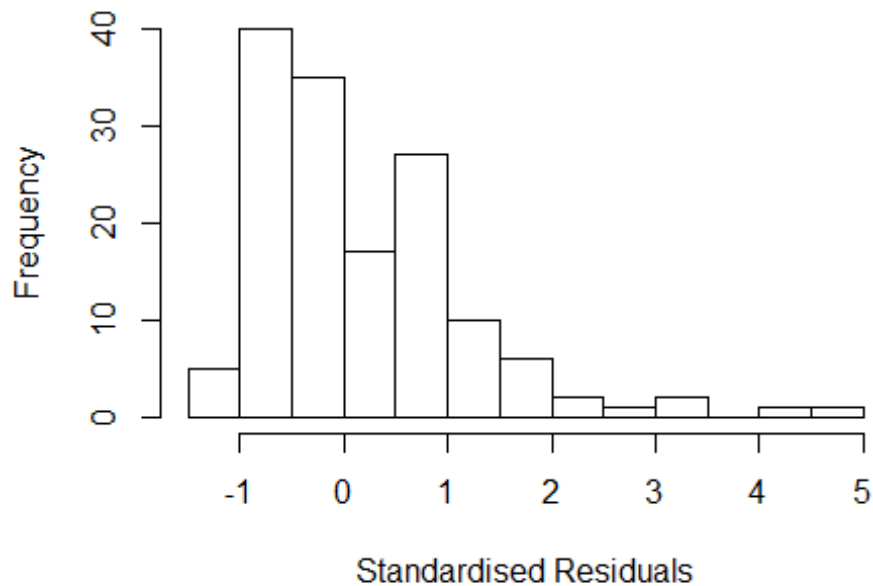
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## Warning in 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 14              NaN
```

Residuals from last author



```
## [1] "List of 5 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 86  77957886942 4.361 2009    1213     1    4.046
## 95  79957891760 3.922 2010    1208     2    3.022
## 130 80053340024 3.290 2011    1208     2    2.545
## 158 84870719154 4.195 2012    1213     1    3.406
## 164 84864343442 5.324 2012    1213     1    4.535
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.422 -0.681 -0.102  0.725  4.539
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.42636    0.16304   8.75 9.2e-15 ***
## LastAuthorFemale1 0.06164    0.16304   0.38 0.70599
## Year1998       -0.89823    0.61272  -1.47 0.14505
## Year1999       -0.30218    0.14556  -2.08 0.03985 *
## Year2001       -1.48800    0.00000  -Inf < 2e-16 ***
## Year2002        0.31270    0.48550   0.64 0.52065
## Year2003       -0.85354    0.13292  -6.42 2.3e-09 ***
## Year2004       -0.00428    0.55019  -0.01 0.99381
```

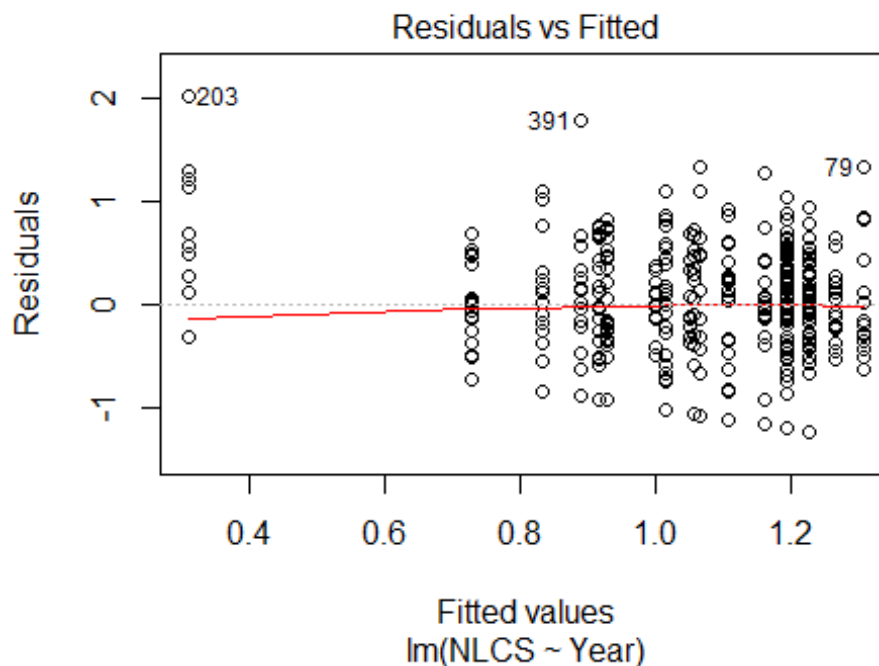
```

## Year2005          0.28720      0.33263      0.86  0.38948
## Year2006          0.08915      0.12534      0.71  0.47817
## Year2007         -0.34942      0.31971     -1.09  0.27644
## Year2008         -0.03568      0.44483     -0.08  0.93618
## Year2009         -1.11512      0.17056     -6.54  1.3e-09 ***
## Year2010         -0.53231      0.23244     -2.29  0.02362 *
## Year2011         -0.74510      0.19105     -3.90  0.00015 ***
## Year2012         -0.64136      0.23262     -2.76  0.00666 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.808
## Multiple R-squared:  0.181, Adjusted R-squared:  0.0874
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 2 observations c(65,132) are outliers with |weight| = 0 ( < 0.00068);
## 8 weights are ~1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0358 0.8820 0.9330 0.8930 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          6.80e-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: 147"
## [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
##   22   29   29   24   67   28   19   24   32   29   24   33   25   40   32
## 2011 2012
##   63   83
##

```



```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 16 14 17 18 42 11 11 19 24 18 16 21 18 26 21
## 2011 2012
## 52 66
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 15 12 11 11 40 10 10 19 21 17 16 17 14 24 20
## 2011 2012
## 47 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 = 26, df = 16, p-value = 0.05
```

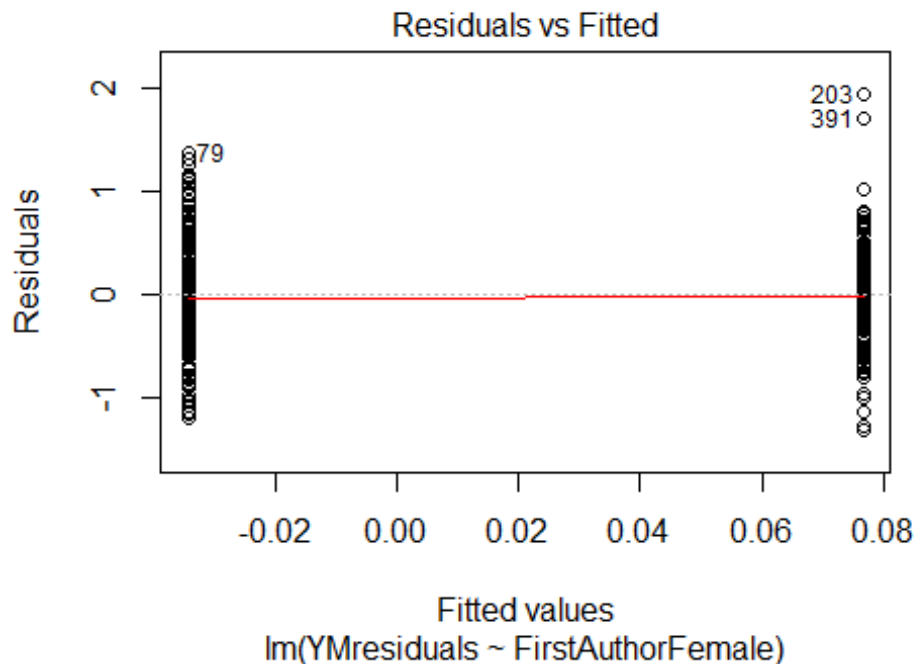


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.041, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 4.56309455714563"
## [1] "Male first author team size 2018 geometric mean: 3.27463488820389"
```

```
## 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 = 520, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.42482725831949"
## [1] "Male last author team size 2018 geometric mean: 3.77059258627322"

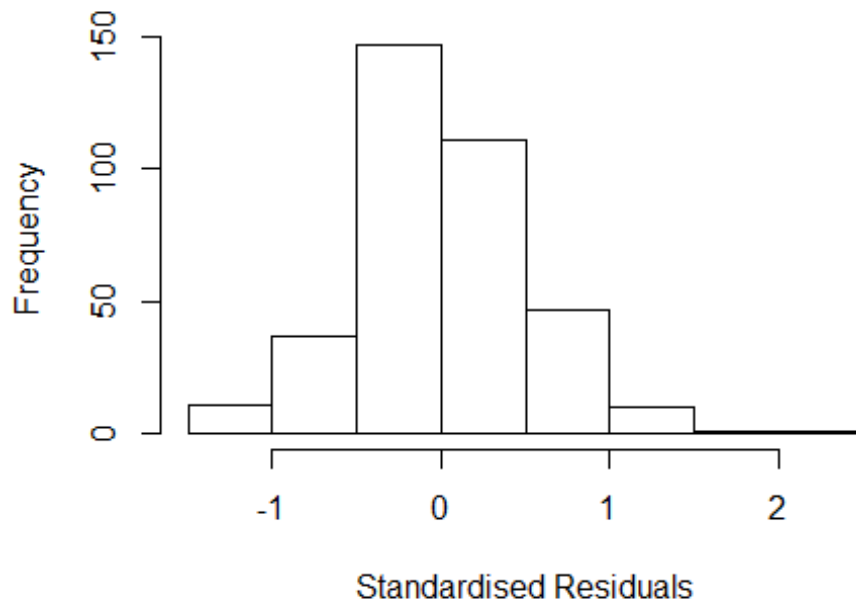
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 400, p-value = 0.2
## 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 1.670 1 1.292
```

## LastAuthorFemale	1.667	1	1.291
## UniqueAuthors	7.820	4	1.293
## Year	10.823	16	1.077

Residuals from first and last author and team size



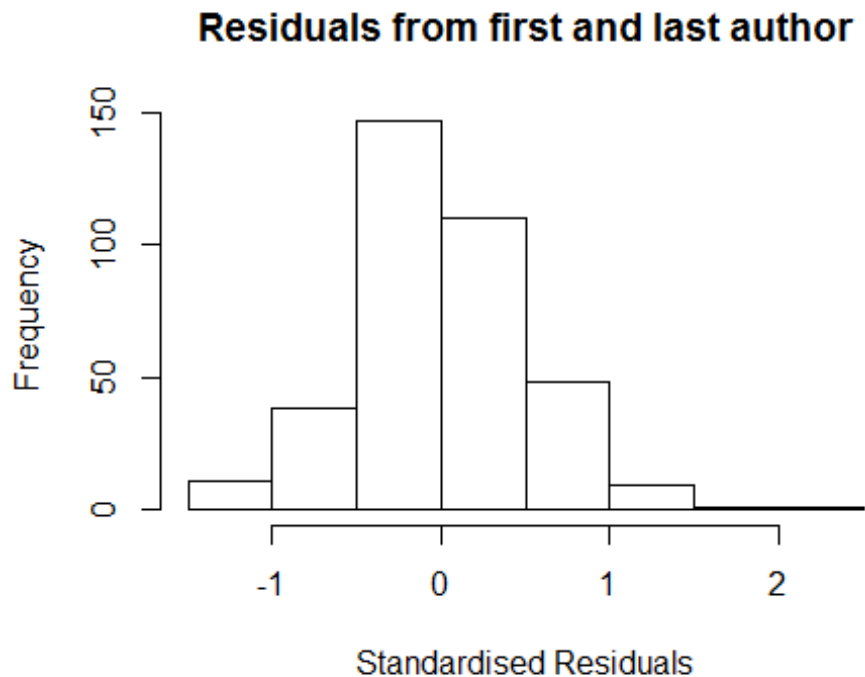
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3308 -0.2577 -0.0359 0.3202 2.2341
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02186 0.20140 5.07 6.4e-07 ***
## FirstAuthorFemale1 0.12342 0.06544 1.89 0.0601 .
## LastAuthorFemale1 -0.15959 0.07357 -2.17 0.0308 *
## UniqueAuthors2 0.13162 0.10917 1.21 0.2288
## UniqueAuthors3 -0.00984 0.09697 -0.10 0.9192
## UniqueAuthors4 0.25626 0.11791 2.17 0.0304 *
```

```

## UniqueAuthors5      0.27920      0.10676      2.62      0.0093 **
## Year1997             0.09030      0.28489      0.32      0.7515
## Year1998             0.10188      0.24371      0.42      0.6762
## Year1999            -0.45842      0.23332     -1.96      0.0503 .
## Year2000            -0.89180      0.21386     -4.17     3.9e-05 ***
## Year2001            -0.15224      0.20327     -0.75      0.4544
## Year2002            -0.09697      0.19336     -0.50      0.6164
## Year2003             0.06339      0.20453      0.31      0.7568
## Year2004            -0.11335      0.22818     -0.50      0.6197
## Year2005            -0.36840      0.20325     -1.81      0.0708 .
## Year2006            -0.29115      0.23214     -1.25      0.2106
## Year2007            -0.14457      0.26914     -0.54      0.5915
## Year2008             0.10089      0.19622      0.51      0.6075
## Year2009            -0.06475      0.22654     -0.29      0.7752
## Year2010            -0.04090      0.22535     -0.18      0.8561
## Year2011             0.02066      0.19334      0.11      0.9149
## Year2012             0.02979      0.19578      0.15      0.8791
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.375, Adjusted R-squared:  0.335
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 88 is an outlier with |weight| = 0 ( < 0.00027);
## 29 weights are ~= 1. The remaining 335 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0016 0.8430 0.9460 0.8850 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      2.74e-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 1.541 1          1.241

```

## LastAuthorFemale	1.634	1	1.278
## Year	1.667	16	1.016



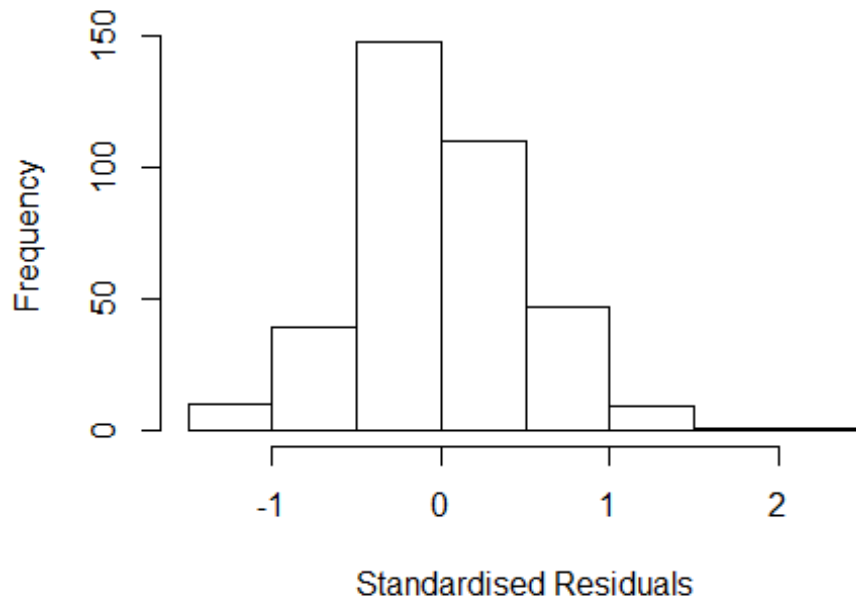
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2446 -0.2873 -0.0535  0.3372  2.2063
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.0973    0.1697   6.47 3.4e-10 ***
## FirstAuthorFemale1  0.1472    0.0673   2.19  0.029 *
## LastAuthorFemale1 -0.1595    0.0762  -2.09  0.037 *
## Year1997         0.0278    0.3034   0.09  0.927
## Year1998         0.1281    0.2407   0.53  0.595
## Year1999        -0.3778    0.2162  -1.75  0.081 .
## Year2000        -0.9633    0.1864  -5.17 4.0e-07 ***
## Year2001        -0.1311    0.1970  -0.67  0.506
## Year2002        -0.1089    0.1938  -0.56  0.575
```

```

## Year2003          0.0726      0.1982      0.37      0.714
## Year2004         -0.1530      0.2163     -0.71      0.480
## Year2005         -0.3651      0.1943     -1.88      0.061 .
## Year2006         -0.3208      0.2210     -1.45      0.148
## Year2007         -0.1444      0.2548     -0.57      0.571
## Year2008          0.1042      0.1908      0.55      0.585
## Year2009         -0.0720      0.2331     -0.31      0.758
## Year2010          0.0143      0.2225      0.06      0.949
## Year2011          0.1009      0.1862      0.54      0.588
## Year2012          0.1195      0.1856      0.64      0.520
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.354, Adjusted R-squared:  0.321
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## observation 88 is an outlier with |weight| = 0 ( < 0.00027);
## 24 weights are ~= 1. The remaining 340 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0044 0.8320 0.9450 0.8770 0.9900 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          2.74e-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 1.243 1          1.115
## Year              1.243 16          1.007

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2968 -0.2867 -0.0422  0.3390  2.1170
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.09946    0.16995   6.47 3.3e-10 ***
## FirstAuthorFemale1 0.10086    0.06237   1.62  0.107
## Year1997        0.04535    0.29728   0.15  0.879
## Year1998        0.14239    0.23965   0.59  0.553
## Year1999       -0.36611    0.21580  -1.70  0.091 .
## Year2000       -0.98930    0.18639  -5.31 2.0e-07 ***
## Year2001       -0.13593    0.19291  -0.70  0.482
## Year2002       -0.13473    0.19082  -0.71  0.481
## Year2003        0.03933    0.19890   0.20  0.843
## Year2004       -0.18543    0.21534  -0.86  0.390
## Year2005       -0.38957    0.19961  -1.95  0.052 .
## Year2006       -0.31807    0.22610  -1.41  0.160
```

```

## Year2007          -0.18131    0.24593   -0.74    0.461
## Year2008           0.07740    0.19175    0.40    0.687
## Year2009          -0.14096    0.23165   -0.61    0.543
## Year2010           0.00155    0.22255    0.01    0.994
## Year2011           0.09470    0.18567    0.51    0.610
## Year2012           0.09651    0.18552    0.52    0.603
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.336, Adjusted R-squared:  0.304
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## observation 88 is an outlier with |weight| = 0 ( < 0.00027);
## 22 weights are ~= 1. The remaining 342 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0561 0.8480 0.9460 0.8840 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.74e-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 1.351 1          1.162
## Year            1.351 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max

```



```

## -1.2661 -0.2945 -0.0582  0.3636  2.2531
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1149    0.1726   6.46 3.6e-10 ***
## LastAuthorFemale1 -0.1055    0.0703  -1.50  0.135
## Year1997          0.0645    0.2792   0.23  0.817
## Year1998          0.1514    0.2402   0.63  0.529
## Year1999         -0.3521    0.2172  -1.62  0.106
## Year2000         -0.9345    0.1884  -4.96 1.1e-06 ***
## Year2001         -0.1134    0.1961  -0.58  0.563
## Year2002         -0.1195    0.1928  -0.62  0.536
## Year2003          0.0699    0.2017   0.35  0.729
## Year2004         -0.1444    0.2179  -0.66  0.508
## Year2005         -0.3353    0.1951  -1.72  0.087 .
## Year2006         -0.3259    0.2267  -1.44  0.151
## Year2007         -0.1328    0.2502  -0.53  0.596
## Year2008          0.1085    0.1934   0.56  0.575
## Year2009         -0.0807    0.2363  -0.34  0.733
## Year2010          0.0159    0.2256   0.07  0.944
## Year2011          0.1353    0.1857   0.73  0.467
## Year2012          0.1513    0.1854   0.82  0.415
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.456
## Multiple R-squared:  0.324, Adjusted R-squared:  0.291
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 88 is an outlier with |weight| = 0 ( < 0.00027);
## 17 weights are ~= 1. The remaining 347 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0187 0.8600 0.9460 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           2.74e-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)

```

```

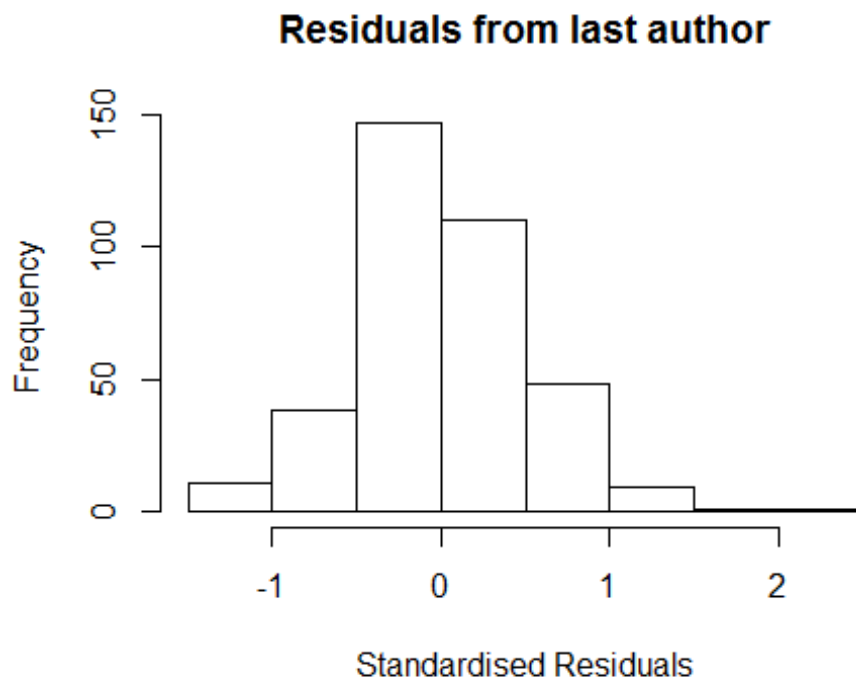
## [1] "Sample size for the above analysis: 365"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1301"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2010
## 2
##
## 2010
## 0
##
## 2010
## 0
## [1] "Heteroscedasticity 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: 8"
## [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: 0"
## [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 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
## 3 2 3 1 2 1 1 2 2 4 5 5 6 3 6
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
## 2 2 2 1 0 1 0 0 1 2 1 4 6 3 5
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
## 2 2 1 1 0 1 0 0 1 1 1 1 4 3 5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.43754385517496"
## [1] "Male first author team size 2018 geometric mean: 8"

## 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, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.82842712474619"
## [1] "Male last author team size 2018 geometric mean: 4.68069463864143"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

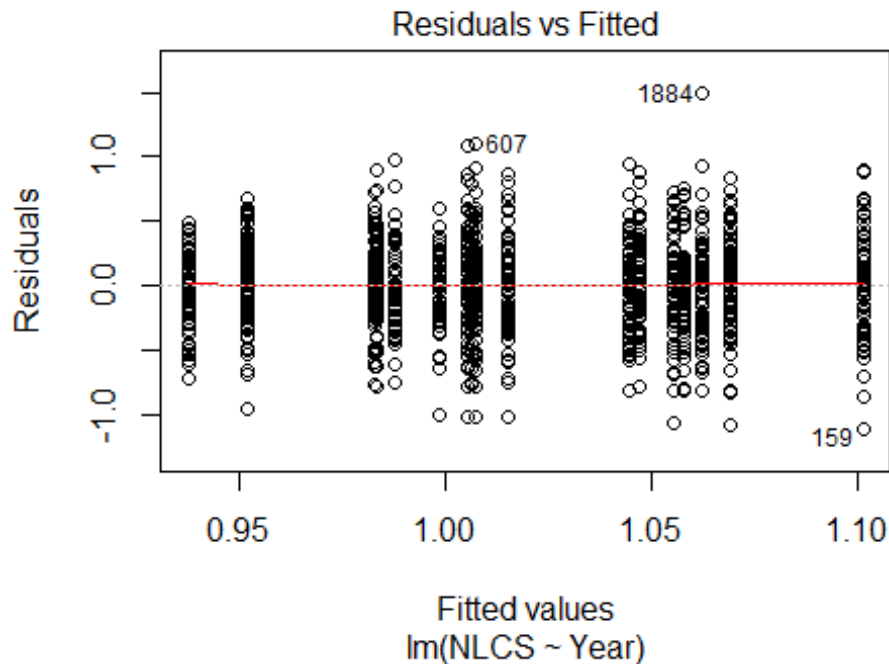


```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1.5, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
## [1] "Regression 2: First author gender, Last author gender, Year as
## [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: 23"
```

```

## [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 2008 2009 2010
## 120 120 103 114 112 108 87 76 86 105 109 102 83 108 107
## 2011 2012
## 111 125
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 76 68 63 70 68 58 54 57 67 80 75 70 62 69 75
## 2011 2012
## 78 97
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 72 65 58 61 57 49 48 47 59 71 66 59 53 61 63
## 2011 2012
## 73 90
## [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 = 16, p-value = 0.002

```



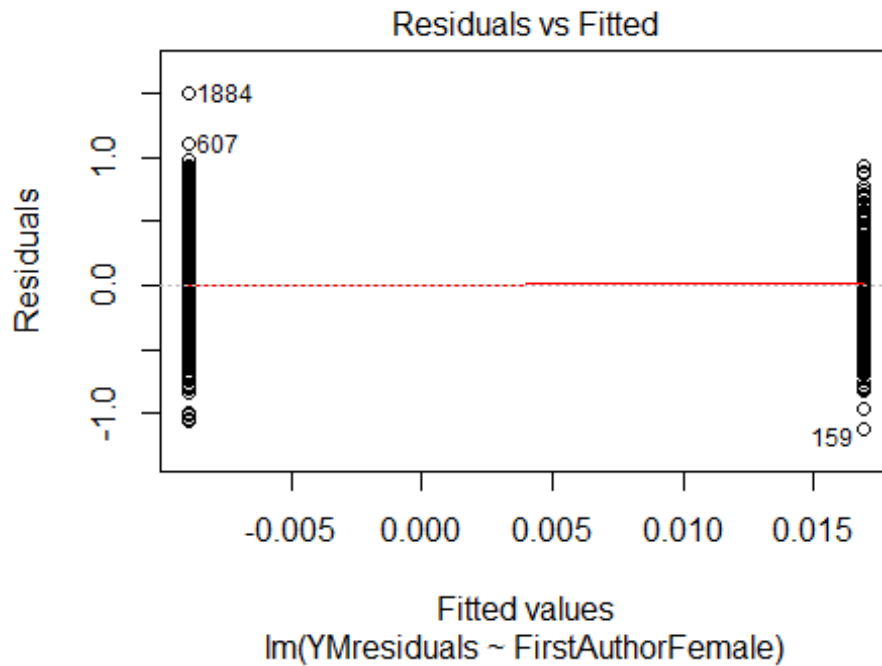
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.5, df = 1, p-value = 0.02

## [1] "Female first author team size 2018 geometric mean: 4.43104806047341"
## [1] "Male first author team size 2018 geometric mean: 4.70310791409192"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

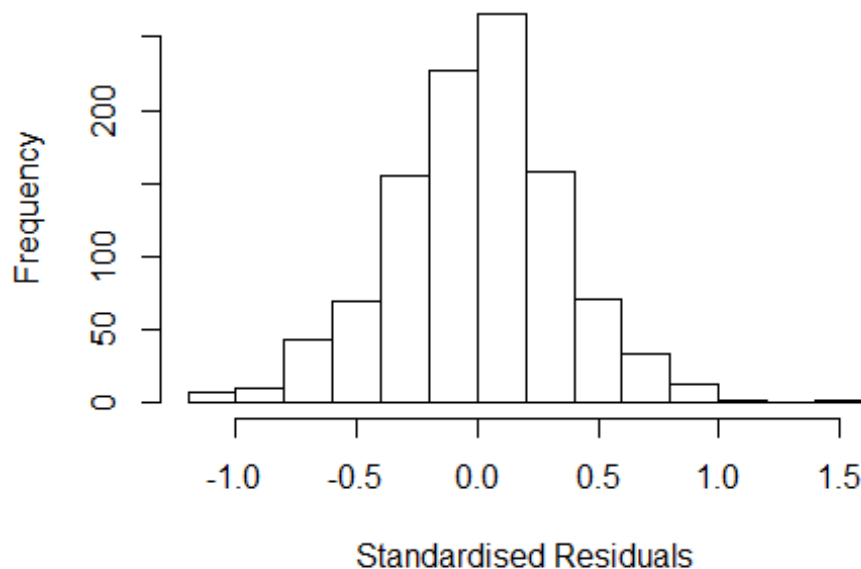
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 260, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.65586659394622"
## [1] "Male last author team size 2018 geometric mean: 4.51248665235337"

## 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 = 310, 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 1.215 1          1.102
## LastAuthorFemale  1.223 1          1.106
## UniqueAuthors    1.555 4          1.057
## Year             1.970 16          1.021
```

Residuals from first and last author and team size



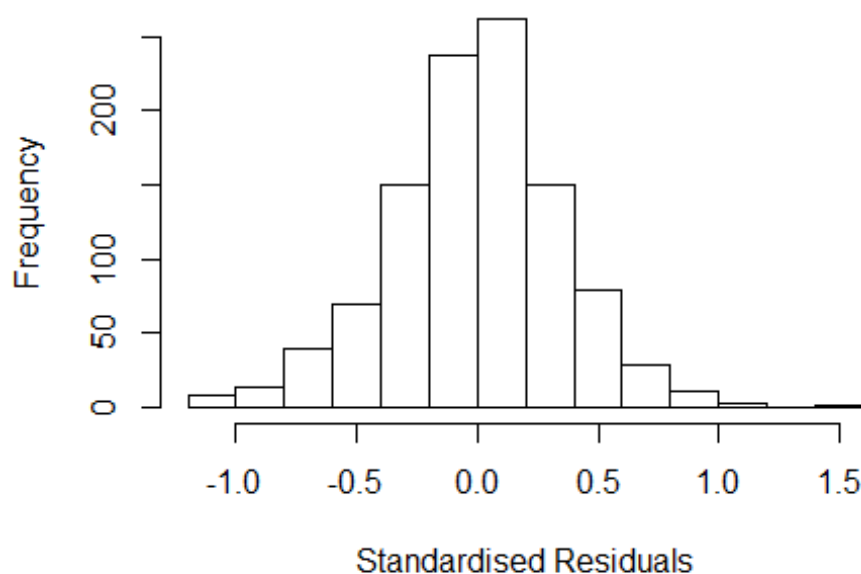
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1066 -0.2220 0.0138 0.2077 1.5989
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.96024 0.05969 16.09 < 2e-16 ***
## FirstAuthorFemale1 0.02338 0.02321 1.01 0.31389
## LastAuthorFemale1 -0.01304 0.02783 -0.47 0.63939
## UniqueAuthors2 0.13083 0.05628 2.32 0.02028 *
## UniqueAuthors3 0.14636 0.05312 2.76 0.00597 **
## UniqueAuthors4 0.19089 0.05532 3.45 0.00058 ***
## UniqueAuthors5 0.28954 0.05487 5.28 1.6e-07 ***
## Year1997 -0.00821 0.07005 -0.12 0.90677
## Year1998 -0.02852 0.07180 -0.40 0.69135
## Year1999 -0.07284 0.05750 -1.27 0.20553
```

```

## Year2000      -0.12527    0.06749   -1.86  0.06373 .
## Year2001      -0.13997    0.06599   -2.12  0.03416 *
## Year2002      -0.13031    0.06180   -2.11  0.03523 *
## Year2003      -0.10622    0.05741   -1.85  0.06459 .
## Year2004      -0.18406    0.05901   -3.12  0.00186 **
## Year2005      -0.16417    0.05716   -2.87  0.00416 **
## Year2006      -0.12719    0.05653   -2.25  0.02467 *
## Year2007      -0.11167    0.06690   -1.67  0.09540 .
## Year2008      -0.12359    0.06517   -1.90  0.05821 .
## Year2009      -0.09896    0.06153   -1.61  0.10809
## Year2010      -0.15246    0.06479   -2.35  0.01880 *
## Year2011      -0.13770    0.06302   -2.18  0.02912 *
## Year2012      -0.19258    0.05725   -3.36  0.00080 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.0656, Adjusted R-squared:  0.0456
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 841 is an outlier with |weight| = 0 ( < 9.5e-05);
## 102 weights are ~= 1. The remaining 949 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.213  0.860  0.952  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           9.51e-05           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01           5.00e-01
##   nResample   max.it   best.r.s   k.fast.s   k.max maxit.scale
##           500         50         2         1         1000         200
##   trace.lev   mts   compute.rd
##           0         1000         0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.138 1 1.067
## LastAuthorFemale 1.226 1 1.107
## Year 1.364 16 1.010

```


Residuals from first and last author



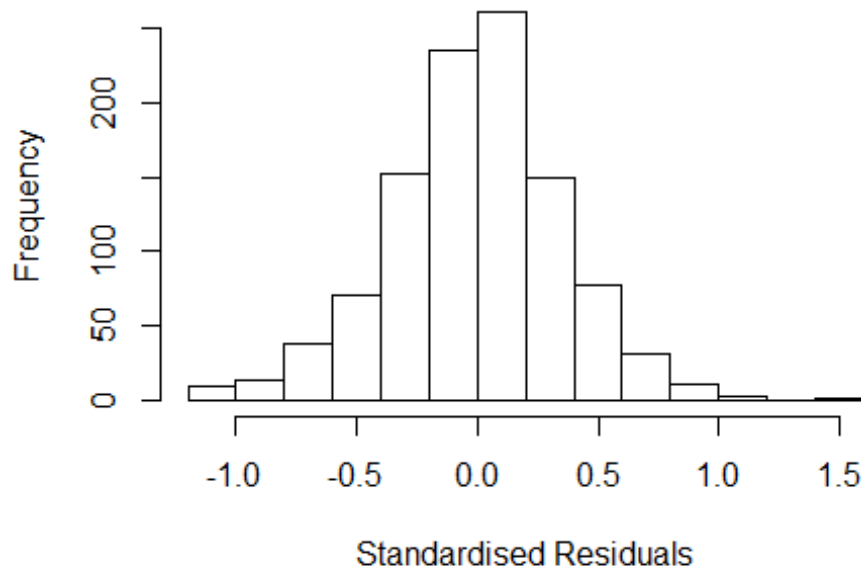
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.15425 -0.22340  0.00864  0.20921  1.55003
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.09394    0.04544   24.08  <2e-16 ***
## FirstAuthorFemale1  0.03901    0.02307    1.69   0.0911 .
## LastAuthorFemale1 -0.01335    0.02809   -0.48   0.6347
## Year1997         0.02130    0.06834    0.31   0.7553
## Year1998        -0.00157    0.07474   -0.02   0.9833
## Year1999        -0.04424    0.05874   -0.75   0.4515
## Year2000        -0.11174    0.07224   -1.55   0.1222
## Year2001        -0.12597    0.06570   -1.92   0.0555 .
## Year2002        -0.10226    0.06327   -1.62   0.1064
## Year2003        -0.07779    0.05813   -1.34   0.1811
## Year2004        -0.15575    0.05959   -2.61   0.0091 **
## Year2005        -0.11612    0.05763   -2.01   0.0442 *
```

```

## Year2006      -0.08626    0.05831   -1.48    0.1394
## Year2007      -0.05284    0.06954   -0.76    0.4475
## Year2008      -0.06442    0.06534   -0.99    0.3244
## Year2009      -0.03846    0.06315   -0.61    0.5426
## Year2010      -0.09097    0.06458   -1.41    0.1593
## Year2011      -0.08127    0.06301   -1.29    0.1974
## Year2012      -0.14163    0.05693   -2.49    0.0130 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.33
## Multiple R-squared:  0.0231, Adjusted R-squared:  0.00603
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## observation 841 is an outlier with |weight| = 0 ( < 9.5e-05);
## 88 weights are ~= 1. The remaining 963 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.194  0.862  0.953  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           9.51e-05           1.82e-12
## warn.limit.reject warn.limit.meanrw
##           5.00e-01                5.00e-01
##   nResample    max.it    best.r.s    k.fast.s    k.max maxit.scale
##           500         50         2         1         1000         200
##   trace.lev    mts    compute.rd
##           0         1000         0
##           psi                subsampling                cov
##           "bisquare"                "nonsingular"                ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.13 1         1.063
## Year              1.13 16         1.004

```

Residuals from first author



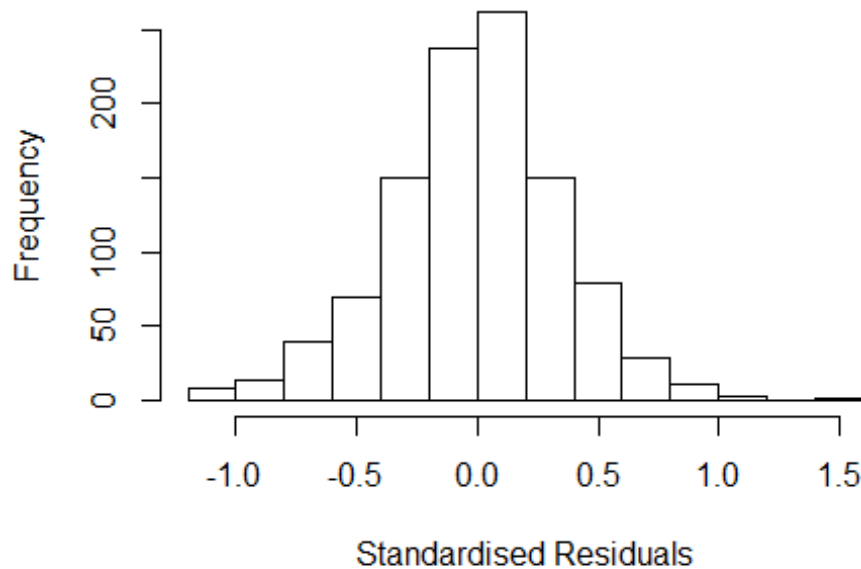
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.15238 -0.22617 0.00808 0.20691 1.55486
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0927 0.0454 24.08 <2e-16 ***
## FirstAuthorFemale1 0.0381 0.0230 1.66 0.0977 .
## Year1997 0.0215 0.0684 0.31 0.7530
## Year1998 -0.0017 0.0748 -0.02 0.9819
## Year1999 -0.0438 0.0588 -0.75 0.4560
## Year2000 -0.1117 0.0723 -1.55 0.1225
## Year2001 -0.1250 0.0657 -1.90 0.0573 .
## Year2002 -0.1028 0.0633 -1.62 0.1047
## Year2003 -0.0792 0.0581 -1.36 0.1732
## Year2004 -0.1571 0.0594 -2.64 0.0083 **
## Year2005 -0.1179 0.0577 -2.04 0.0413 *
## Year2006 -0.0870 0.0583 -1.49 0.1360
```

```

## Year2007          -0.0562      0.0688   -0.82   0.4137
## Year2008          -0.0670      0.0646   -1.04   0.2998
## Year2009          -0.0411      0.0627   -0.66   0.5117
## Year2010          -0.0946      0.0636   -1.49   0.1374
## Year2011          -0.0852      0.0623   -1.37   0.1719
## Year2012          -0.1449      0.0566   -2.56   0.0106 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.329
## Multiple R-squared:  0.0229, Adjusted R-squared:  0.0068
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## observation 841 is an outlier with |weight| = 0 ( < 9.5e-05);
## 90 weights are ~= 1. The remaining 961 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.195  0.862  0.952  0.894  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      9.51e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.207 1          1.099
## Year          1.207 16          1.006

```

Residuals from last author



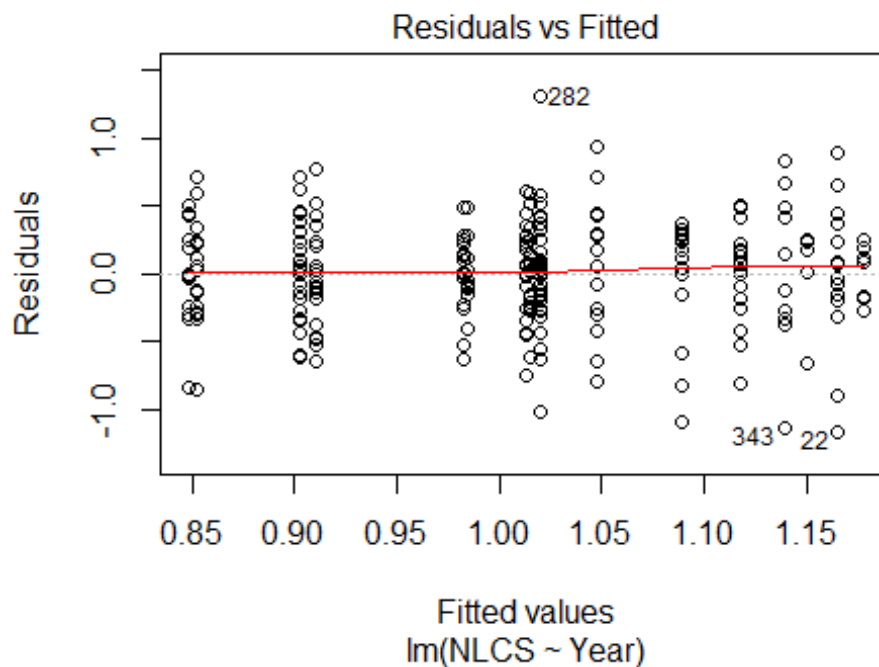
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.12123 -0.21965  0.00747  0.21020  1.53156
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.10719    0.04431   24.99  <2e-16 ***
## LastAuthorFemale1 -0.00980    0.02787   -0.35  0.7252
## Year1997         0.01403    0.06835    0.21  0.8374
## Year1998        -0.00682    0.07418   -0.09  0.9268
## Year1999        -0.04654    0.05900   -0.79  0.4304
## Year2000        -0.11044    0.07226   -1.53  0.1268
## Year2001        -0.12994    0.06517   -1.99  0.0464 *
## Year2002        -0.10364    0.06409   -1.62  0.1062
## Year2003        -0.07586    0.05861   -1.29  0.1958
## Year2004        -0.15633    0.05997   -2.61  0.0093 **
## Year2005        -0.11425    0.05766   -1.98  0.0478 *
## Year2006        -0.08647    0.05885   -1.47  0.1420
```

```

## Year2007          -0.05458      0.06895    -0.79    0.4288
## Year2008          -0.06263      0.06567    -0.95    0.3405
## Year2009          -0.03819      0.06325    -0.60    0.5461
## Year2010          -0.08575      0.06516    -1.32    0.1885
## Year2011          -0.08034      0.06352    -1.26    0.2062
## Year2012          -0.13745      0.05720    -2.40    0.0164 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.33
## Multiple R-squared:  0.0199, Adjusted R-squared:  0.00383
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 92 weights are ~= 1. The remaining 960 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0005 0.8630 0.9530 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      9.51e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 1052"
## [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
##   24   21   27   21   20   15   10   24   22   30   35   23   23   32   29
## 2011 2012
##   35   34
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   17   15   12   11   15    5    7   15   18   18   16   12   15   17   22
## 2011 2012

```

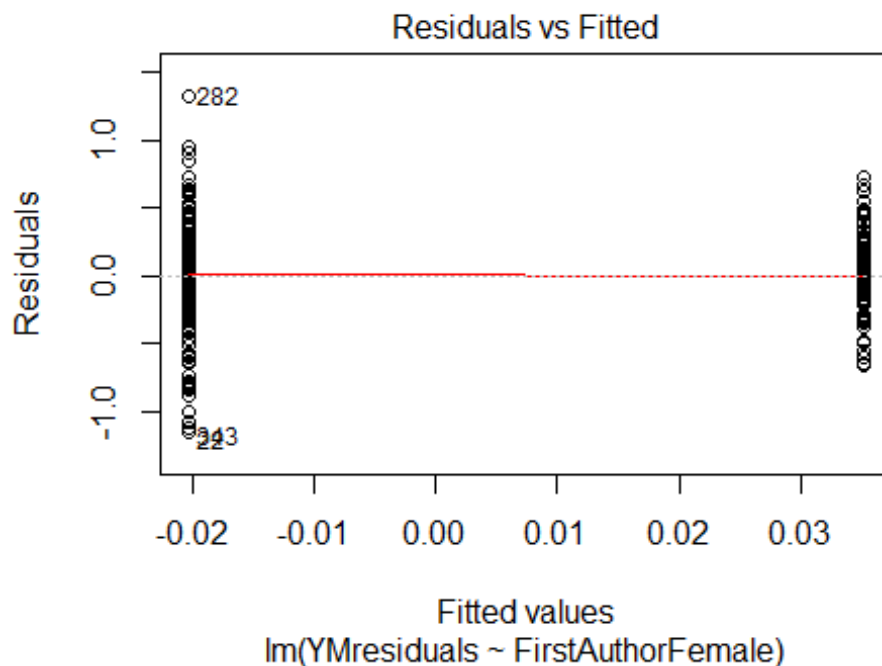
```
## 19 22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 15 15 11 10 14 4 5 11 16 12 15 10 12 16 19
## 2011 2012
## 17 20
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 25, df = 16, p-value = 0.08
```



```
##
## 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: 5.06101774236755"
## [1] "Male first author team size 2018 geometric mean: 3.41347367369015"
## 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 = 67, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 4.20782538349378"

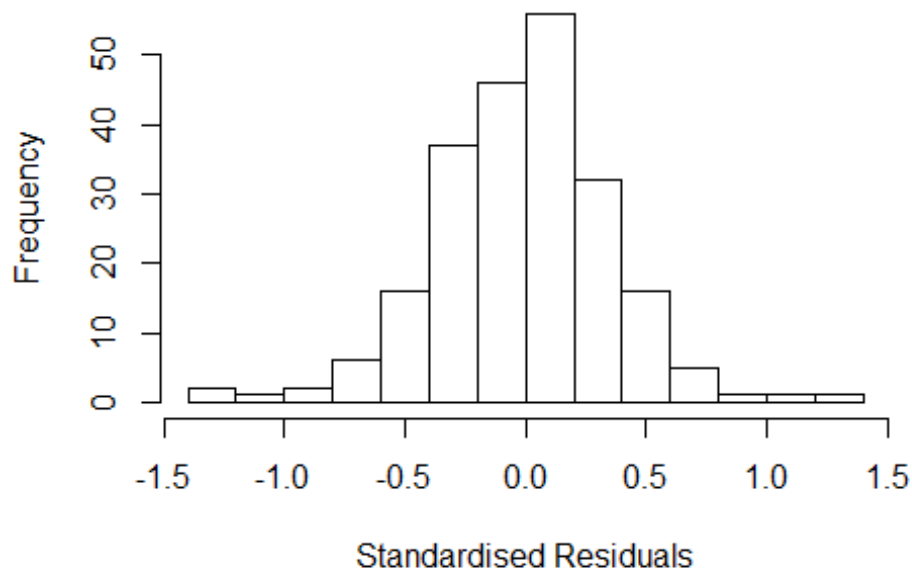
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, p-value = 0.8
## 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	1.931	1	1.390
LastAuthorFemale	1.878	1	1.370
UniqueAuthors	6.490	4	1.263
Year	12.436	16	1.082

Residuals from first and last author and team size



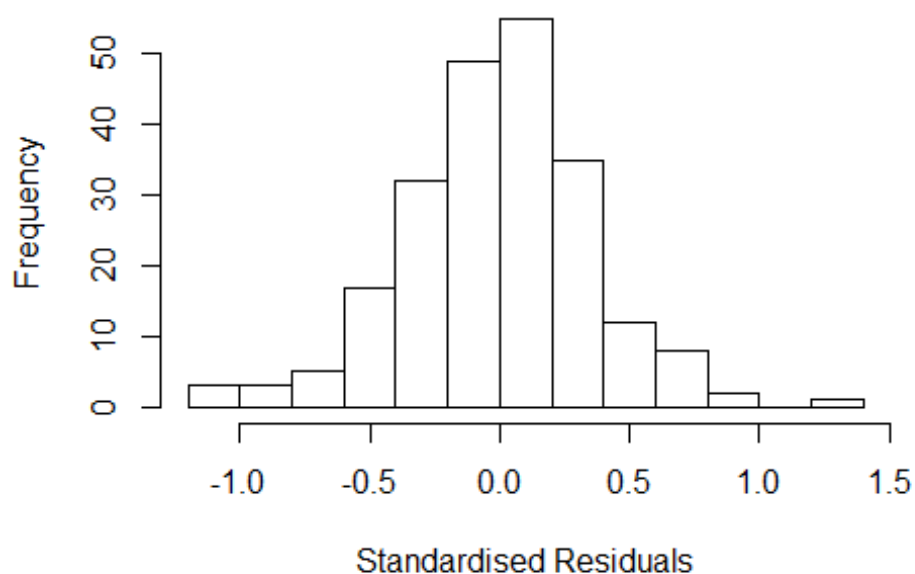
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.21984 -0.21994 0.00528 0.20526 1.25156
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06179 0.26367 4.03 8e-05 ***
## FirstAuthorFemale1 0.07565 0.05145 1.47 0.143
## LastAuthorFemale1 -0.11601 0.07726 -1.50 0.135
## UniqueAuthors2 0.15805 0.18938 0.83 0.405
## UniqueAuthors3 0.12014 0.18845 0.64 0.525
## UniqueAuthors4 0.19086 0.18680 1.02 0.308
## UniqueAuthors5 0.09347 0.18619 0.50 0.616
## Year1997 -0.15370 0.21387 -0.72 0.473
## Year1998 -0.25590 0.18154 -1.41 0.160
## Year1999 -0.25783 0.15323 -1.68 0.094 .
```

```

## Year2000      -0.26267    0.14117   -1.86    0.064 .
## Year2001      0.09746    0.14473    0.67    0.501
## Year2002     -0.02070    0.16847   -0.12    0.902
## Year2003     -0.08172    0.15403   -0.53    0.596
## Year2004     -0.14710    0.14885   -0.99    0.324
## Year2005     -0.10949    0.16083   -0.68    0.497
## Year2006     -0.00635    0.15887   -0.04    0.968
## Year2007      0.05432    0.24268    0.22    0.823
## Year2008     -0.37021    0.15968   -2.32    0.021 *
## Year2009     -0.09417    0.16434   -0.57    0.567
## Year2010     -0.31254    0.14760   -2.12    0.035 *
## Year2011     -0.10815    0.14714   -0.74    0.463
## Year2012     -0.29571    0.15704   -1.88    0.061 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.334
## Multiple R-squared:  0.15,   Adjusted R-squared:  0.056
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 203 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.131  0.873  0.955  0.895  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      4.50e-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 1.923 1      1.387
## LastAuthorFemale  1.452 1      1.205
## Year              2.549 16      1.030

```

Residuals from first and last author



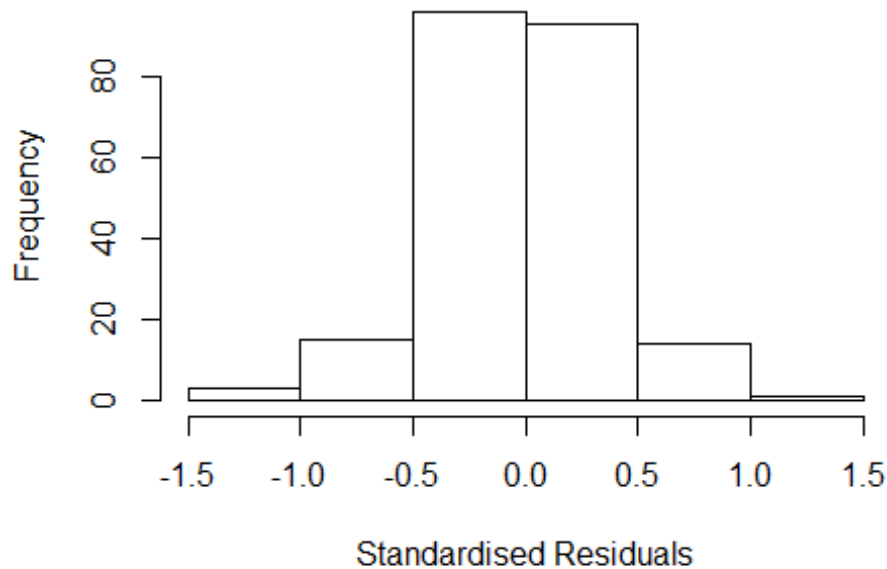
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1920 -0.2321 0.0073 0.2073 1.2218
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.19204 0.14035 8.49 4.3e-15 ***
## FirstAuthorFemale1 0.08193 0.05228 1.57 0.119
## LastAuthorFemale1 -0.12967 0.07240 -1.79 0.075 .
## Year1997 -0.14387 0.21205 -0.68 0.498
## Year1998 -0.28261 0.16494 -1.71 0.088 .
## Year1999 -0.23856 0.15269 -1.56 0.120
## Year2000 -0.24376 0.14446 -1.69 0.093 .
## Year2001 0.12407 0.14805 0.84 0.403
## Year2002 -0.00513 0.17487 -0.03 0.977
## Year2003 -0.07996 0.15649 -0.51 0.610
## Year2004 -0.15659 0.14444 -1.08 0.280
## Year2005 -0.08980 0.16944 -0.53 0.597
```

```

## Year2006      -0.02332    0.15910   -0.15    0.884
## Year2007      0.06881    0.23857    0.29    0.773
## Year2008     -0.35207    0.16726   -2.10    0.037 *
## Year2009     -0.11379    0.15860   -0.72    0.474
## Year2010     -0.32010    0.14925   -2.14    0.033 *
## Year2011     -0.09176    0.14736   -0.62    0.534
## Year2012     -0.30085    0.16239   -1.85    0.065 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.135, Adjusted R-squared:  0.0586
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 213 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.164  0.889  0.957  0.901  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      4.50e-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 1.846 1      1.359
## Year              1.846 16      1.019

```

Residuals from first author



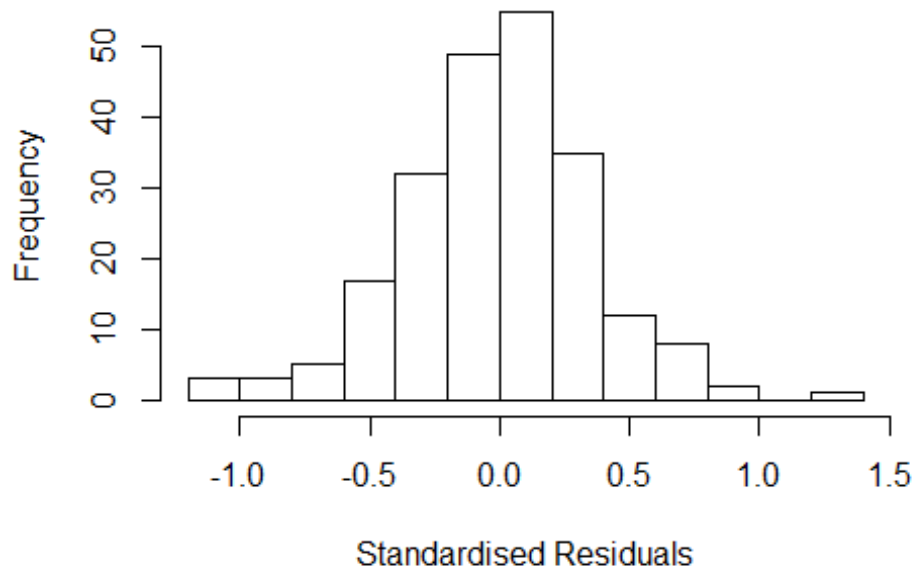
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2869 -0.2247 -0.0109 0.2100 1.2538
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1722 0.1348 8.69 1.2e-15 ***
## FirstAuthorFemale1 0.0681 0.0519 1.31 0.191
## Year1997 -0.1256 0.2223 -0.57 0.573
## Year1998 -0.2602 0.1606 -1.62 0.107
## Year1999 -0.2129 0.1464 -1.45 0.147
## Year2000 -0.2399 0.1424 -1.68 0.094 .
## Year2001 0.1439 0.1428 1.01 0.315
## Year2002 0.0171 0.1691 0.10 0.920
## Year2003 -0.0958 0.1567 -0.61 0.542
## Year2004 -0.1455 0.1409 -1.03 0.303
## Year2005 -0.1020 0.1661 -0.61 0.540
## Year2006 -0.0167 0.1540 -0.11 0.914
```

```

## Year2007          0.1147      0.2294      0.50      0.618
## Year2008         -0.3513      0.1597     -2.20      0.029 *
## Year2009         -0.1207      0.1558     -0.77      0.439
## Year2010         -0.3183      0.1445     -2.20      0.029 *
## Year2011         -0.1121      0.1448     -0.77      0.440
## Year2012         -0.3019      0.1564     -1.93      0.055 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.332
## Multiple R-squared:  0.126, Adjusted R-squared:  0.0531
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.101  0.875   0.950   0.889   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      4.50e-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 1.442 1      1.201
## Year              1.442 16      1.012

```

Residuals from last author



```
## [1] "List of 0 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
## -1.23234 -0.24940 0.00972 0.21477 1.17047
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2323 0.1347 9.15 <2e-16 ***
## LastAuthorFemale1 -0.1167 0.0717 -1.63 0.105
## Year1997 -0.1774 0.2092 -0.85 0.397
## Year1998 -0.3079 0.1666 -1.85 0.066 .
## Year1999 -0.2422 0.1542 -1.57 0.118
## Year2000 -0.2463 0.1481 -1.66 0.098 .
## Year2001 0.0837 0.1427 0.59 0.558
## Year2002 -0.0319 0.1646 -0.19 0.847
## Year2003 -0.0928 0.1625 -0.57 0.568
## Year2004 -0.1481 0.1477 -1.00 0.317
## Year2005 -0.0788 0.1733 -0.45 0.650
## Year2006 -0.0349 0.1615 -0.22 0.829
```

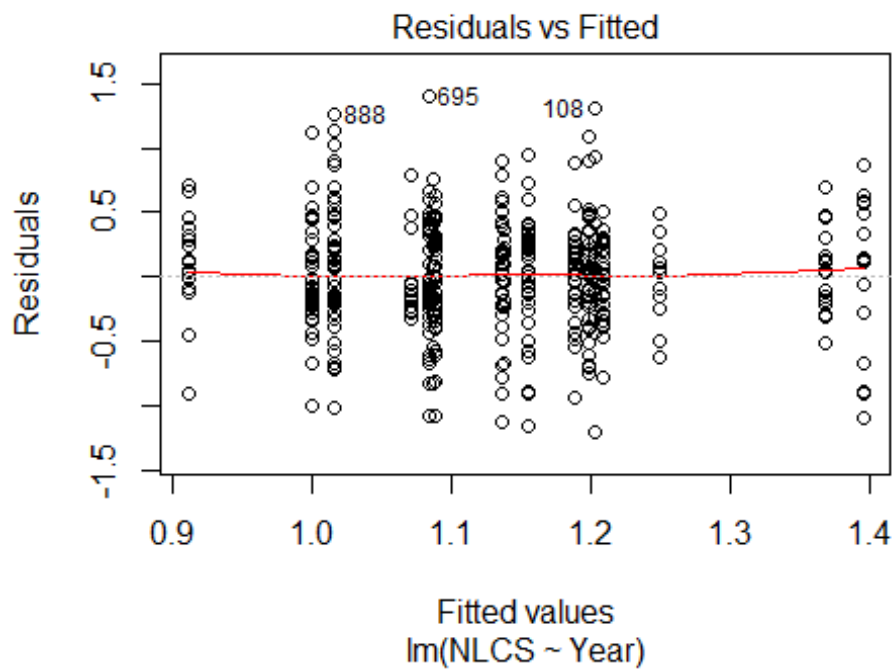
```

## Year2007          0.0468      0.2361      0.20      0.843
## Year2008          -0.3574      0.1641     -2.18      0.031 *
## Year2009          -0.1456      0.1559     -0.93      0.351
## Year2010          -0.3254      0.1521     -2.14      0.034 *
## Year2011          -0.0786      0.1521     -0.52      0.606
## Year2012          -0.3083      0.1646     -1.87      0.063 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.347
## Multiple R-squared:  0.126, Adjusted R-squared:  0.053
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.181  0.869  0.948  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      4.50e-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: 222"
## [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
##   45   50   49   29   43   36   47   26   28   46   47   59   54   51   55
## 2011 2012
##   68   62
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   18   17   19   18   12   12   28   15   10   27   28   33   36   34   40
## 2011 2012

```



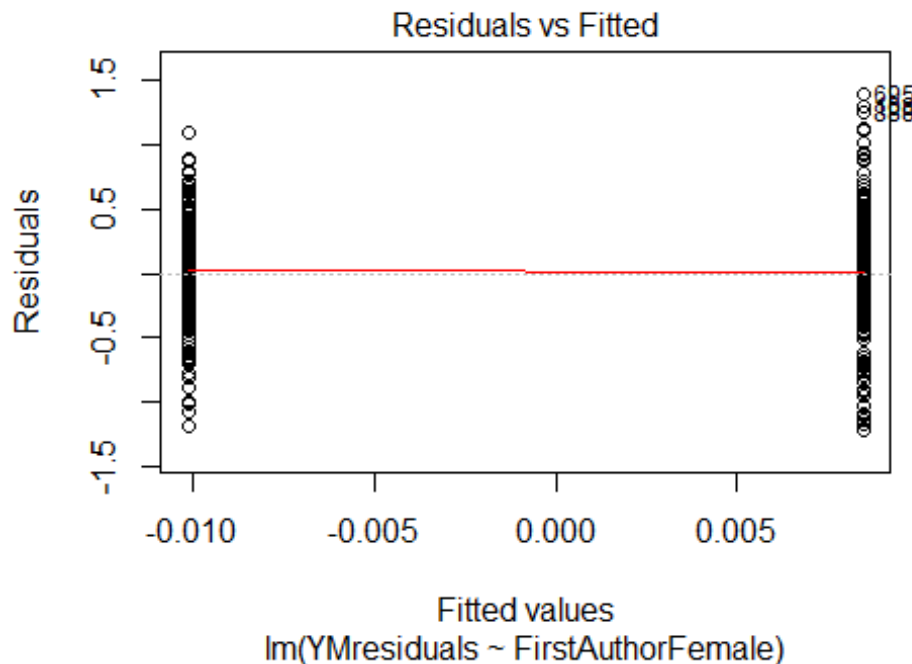
```
## 50 42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 17 17 17 17 10 11 22 12 9 20 24 31 31 29 37
## 2011 2012
## 45 33
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 3, df = 1, p-value = 0.08
## [1] "Female first author team size 2018 geometric mean: 2.8513481795695"
## [1] "Male first author team size 2018 geometric mean: 3.84958105051036"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

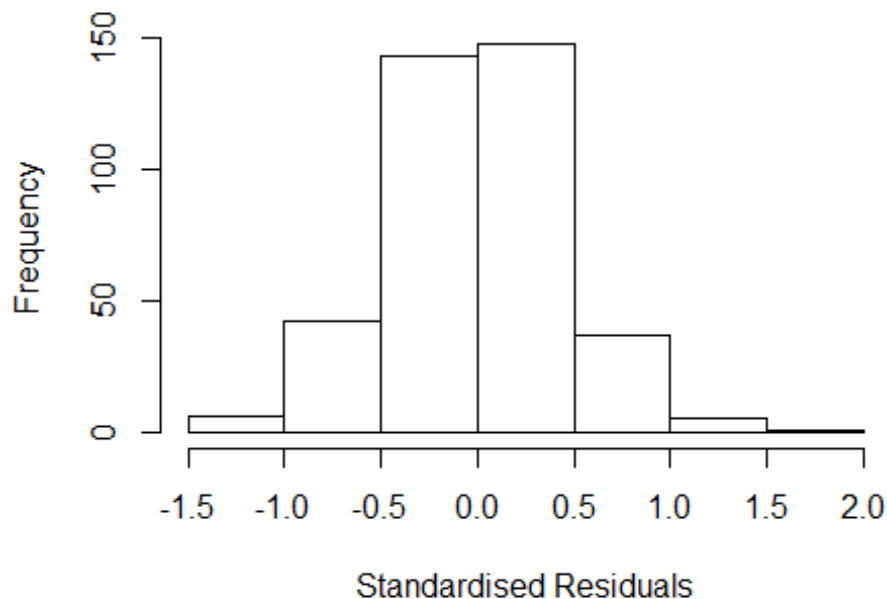
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.1160564564634"
## [1] "Male last author team size 2018 geometric mean: 3.97265525644976"

## 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 = 73, 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 1.251  1      1.118
## LastAuthorFemale  1.266  1      1.125
## UniqueAuthors    2.957  4      1.145
## Year              3.729 16      1.042
```

Residuals from first and last author and team size



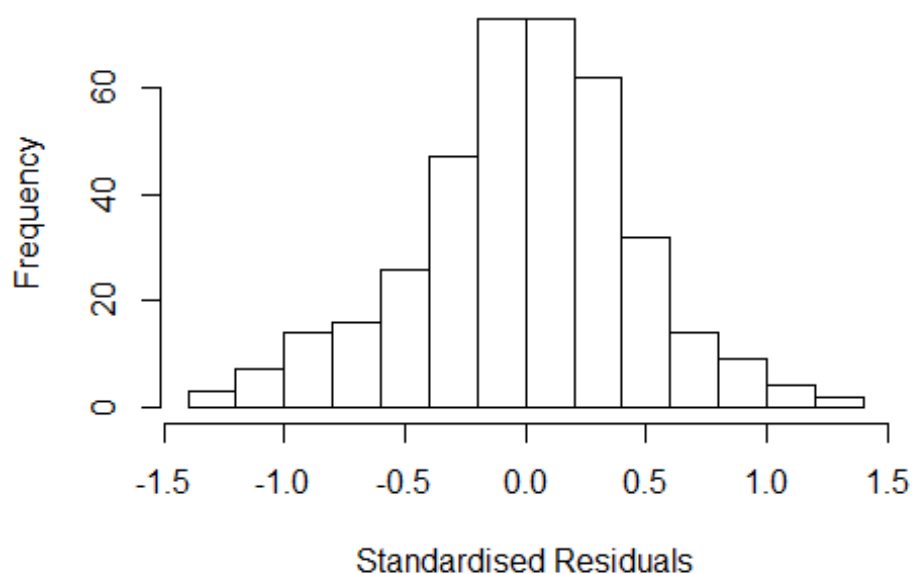
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.35940 -0.25836 0.00201 0.28597 1.57346
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4179 0.2042 6.94 1.8e-11 ***
## FirstAuthorFemale1 -0.0145 0.0495 -0.29 0.7694
## LastAuthorFemale1 -0.0712 0.0613 -1.16 0.2464
## UniqueAuthors2 0.0903 0.1311 0.69 0.4916
## UniqueAuthors3 0.1347 0.1210 1.11 0.2665
## UniqueAuthors4 0.2636 0.1148 2.30 0.0223 *
## UniqueAuthors5 0.2387 0.1134 2.11 0.0359 *
## Year1997 -0.3221 0.2210 -1.46 0.1458
## Year1998 -0.2431 0.1847 -1.32 0.1889
## Year1999 -0.6360 0.2091 -3.04 0.0025 **
```

```

## Year2000          -0.4295      0.2030    -2.12    0.0351 *
## Year2001          -0.4531      0.2273    -1.99    0.0470 *
## Year2002          -0.4450      0.2156    -2.06    0.0397 *
## Year2003          -0.3425      0.1977    -1.73    0.0840 .
## Year2004          -0.3617      0.1770    -2.04    0.0418 *
## Year2005          -0.3507      0.1870    -1.88    0.0616 .
## Year2006          -0.3766      0.1753    -2.15    0.0324 *
## Year2007          -0.4156      0.1873    -2.22    0.0271 *
## Year2008          -0.5054      0.1899    -2.66    0.0081 **
## Year2009          -0.4461      0.1874    -2.38    0.0178 *
## Year2010          -0.4759      0.1792    -2.65    0.0083 **
## Year2011          -0.5473      0.1956    -2.80    0.0054 **
## Year2012          -0.5019      0.1851    -2.71    0.0070 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.118, Adjusted R-squared:  0.0637
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 353 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.120  0.868  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      2.62e-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 1.171 1      1.082
## LastAuthorFemale  1.216 1      1.103
## Year              1.404 16      1.011

```

Residuals from first and last author



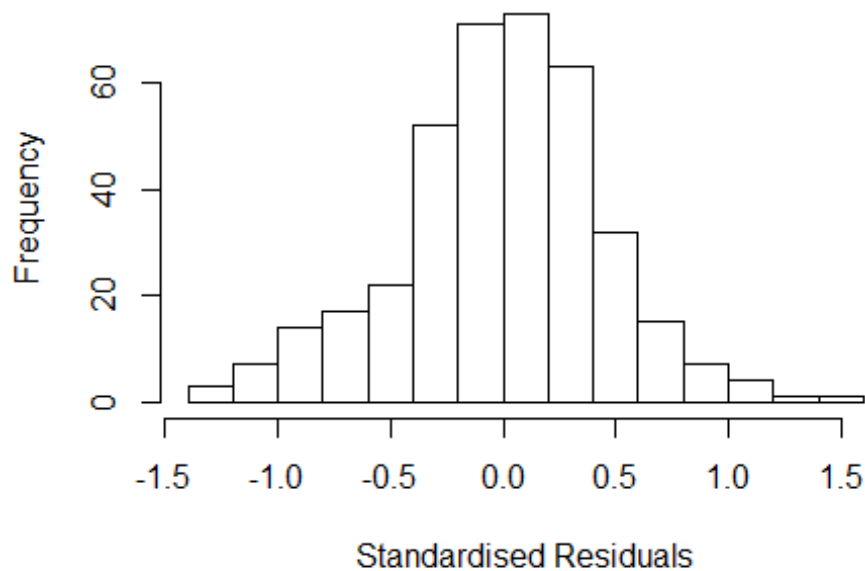
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.26996 -0.26437 0.00823 0.27066 1.39627
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.572964 0.171659 9.16 <2e-16 ***
## FirstAuthorFemale1 -0.000123 0.048423 0.00 0.9980
## LastAuthorFemale1 -0.072386 0.061959 -1.17 0.2435
## Year1997 -0.335899 0.214709 -1.56 0.1186
## Year1998 -0.232910 0.190000 -1.23 0.2211
## Year1999 -0.617592 0.212690 -2.90 0.0039 **
## Year2000 -0.478803 0.210651 -2.27 0.0236 *
## Year2001 -0.457569 0.224545 -2.04 0.0423 *
## Year2002 -0.416054 0.219033 -1.90 0.0583 .
## Year2003 -0.304191 0.199788 -1.52 0.1287
## Year2004 -0.352669 0.183178 -1.93 0.0550 .
## Year2005 -0.350173 0.194270 -1.80 0.0723 .
```

```

## Year2006      -0.355619    0.182123    -1.95    0.0516 .
## Year2007      -0.374207    0.193392    -1.93    0.0538 .
## Year2008      -0.483234    0.192290    -2.51    0.0124 *
## Year2009      -0.422673    0.190893    -2.21    0.0274 *
## Year2010      -0.430675    0.184203    -2.34    0.0199 *
## Year2011      -0.564281    0.195672    -2.88    0.0042 **
## Year2012      -0.511466    0.187135    -2.73    0.0066 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.0844, Adjusted R-squared:  0.039
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 352 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.244  0.869  0.955  0.894  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      2.62e-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 1.166 1      1.080
## Year              1.166 16      1.005

```

Residuals from first author



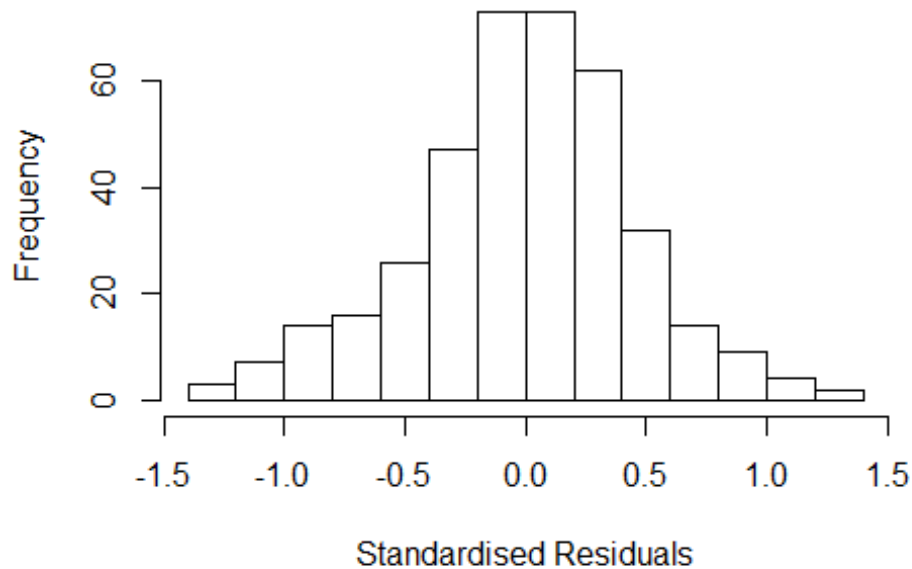
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2754 -0.2502 0.0143 0.2713 1.4100
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.57835 0.17303 9.12 <2e-16 ***
## FirstAuthorFemale1 -0.00601 0.04823 -0.12 0.9010
## Year1997 -0.35236 0.21236 -1.66 0.0979 .
## Year1998 -0.25080 0.18875 -1.33 0.1848
## Year1999 -0.64202 0.21017 -3.05 0.0024 **
## Year2000 -0.49046 0.21076 -2.33 0.0205 *
## Year2001 -0.46205 0.22384 -2.06 0.0397 *
## Year2002 -0.44371 0.21528 -2.06 0.0400 *
## Year2003 -0.33527 0.19536 -1.72 0.0870 .
## Year2004 -0.36330 0.18499 -1.96 0.0503 .
## Year2005 -0.36842 0.19478 -1.89 0.0594 .
## Year2006 -0.36160 0.18306 -1.98 0.0490 *
```

```

## Year2007          -0.39922    0.19206   -2.08    0.0384 *
## Year2008          -0.50240    0.19113   -2.63    0.0089 **
## Year2009          -0.43159    0.19153   -2.25    0.0248 *
## Year2010          -0.44891    0.18374   -2.44    0.0150 *
## Year2011          -0.59555    0.19552   -3.05    0.0025 **
## Year2012          -0.52823    0.18704   -2.82    0.0050 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.418
## Multiple R-squared:  0.0815, Adjusted R-squared:  0.0386
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 354 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.232  0.871  0.955  0.894  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      2.62e-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 1.206 1          1.098
## Year            1.206 16          1.006

```


Residuals from last author



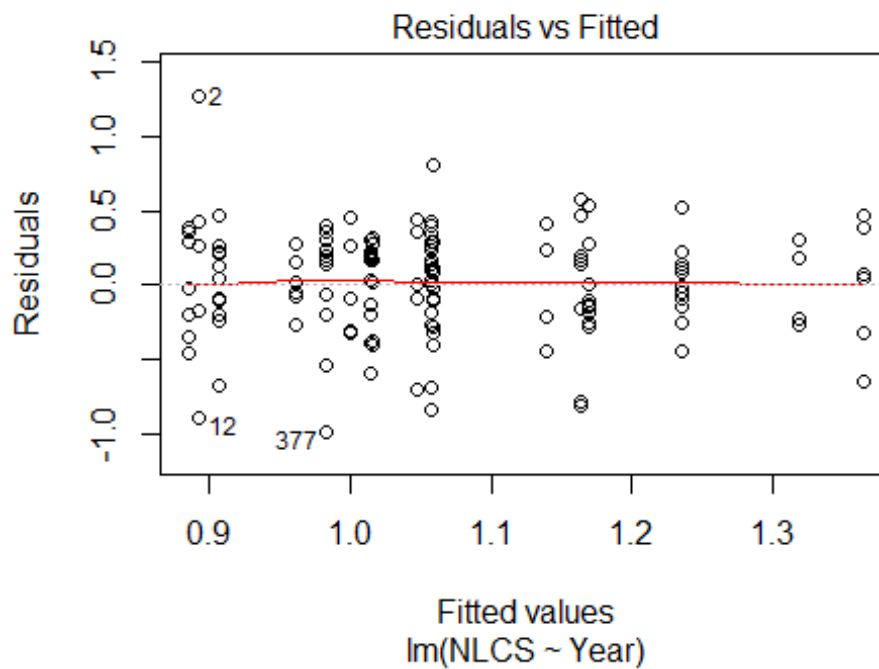
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.26886 -0.26433 0.00851 0.27123 1.39623
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5719 0.1706 9.21 <2e-16 ***
## LastAuthorFemale1 -0.0726 0.0617 -1.18 0.2402
## Year1997 -0.3350 0.2146 -1.56 0.1193
## Year1998 -0.2317 0.1894 -1.22 0.2221
## Year1999 -0.6169 0.2123 -2.91 0.0039 **
## Year2000 -0.4775 0.2102 -2.27 0.0237 *
## Year2001 -0.4566 0.2242 -2.04 0.0424 *
## Year2002 -0.4152 0.2181 -1.90 0.0578 .
## Year2003 -0.3031 0.1993 -1.52 0.1292
## Year2004 -0.3516 0.1828 -1.92 0.0553 .
## Year2005 -0.3491 0.1940 -1.80 0.0727 .
## Year2006 -0.3546 0.1816 -1.95 0.0516 .
```

```

## Year2007          -0.3729      0.1931    -1.93    0.0542 .
## Year2008          -0.4821      0.1917    -2.51    0.0124 *
## Year2009          -0.4218      0.1905    -2.21    0.0274 *
## Year2010          -0.4298      0.1837    -2.34    0.0198 *
## Year2011          -0.5630      0.1954    -2.88    0.0042 **
## Year2012          -0.5104      0.1867    -2.73    0.0066 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.0841, Adjusted R-squared:  0.0414
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 352 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.248  0.870  0.955  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      2.62e-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: 382"
## [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
##   10   27   11   26    9   21   15   11   15    8   13   13   16   11   10
## 2011 2012
##   16   18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6   12    6   18    6    5    9    4   11    5   11   10   11    4    7
## 2011 2012

```

```
##      7    11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6    12     4    17     5     5     9     4     9     5    10    10    11     3     7
## 2011 2012
##      7    10
## [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 = 0.08
```



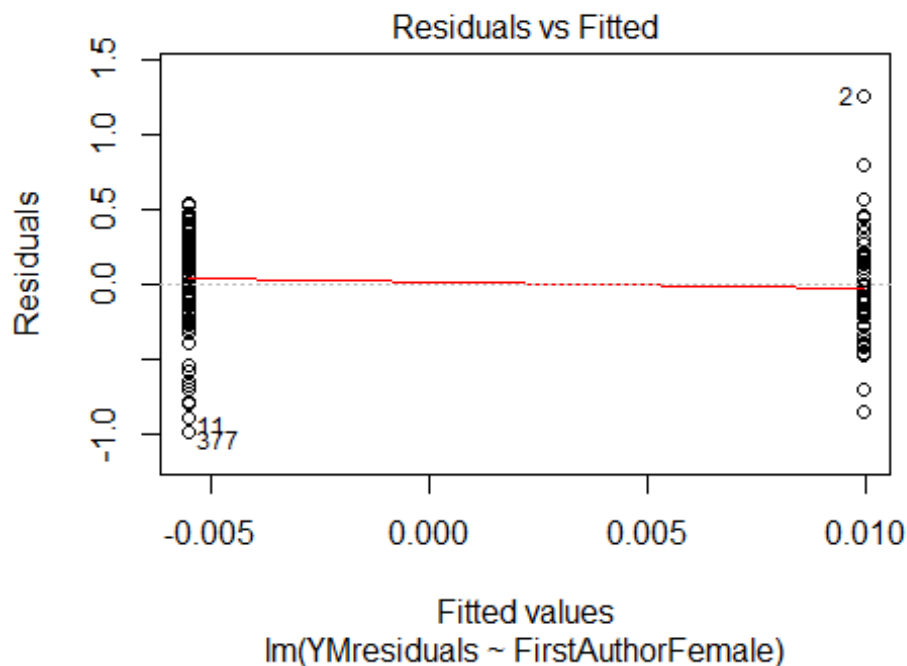
```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.29, df = 1, p-value = 0.6

## [1] "Female first author team size 2018 geometric mean: 4.57885697021333"
## [1] "Male first author team size 2018 geometric mean: 6.49068444639842"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

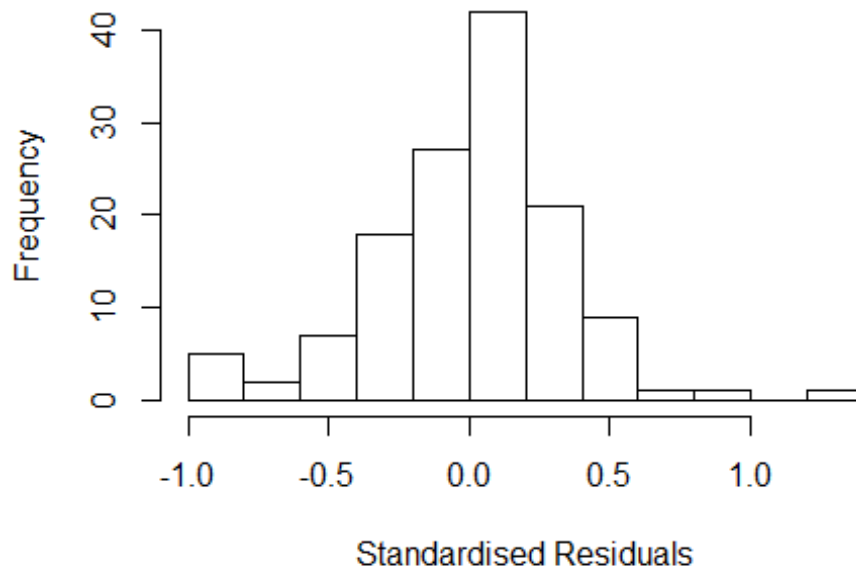
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5.5, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.26429605180997"
## [1] "Male last author team size 2018 geometric mean: 6.16014057648205"

## 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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"
##
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.171 1      1.781
## LastAuthorFemale  1.545 1      1.243
## UniqueAuthors    14.327 4      1.395
## Year              38.766 16     1.121
```

Residuals from first and last author and team size



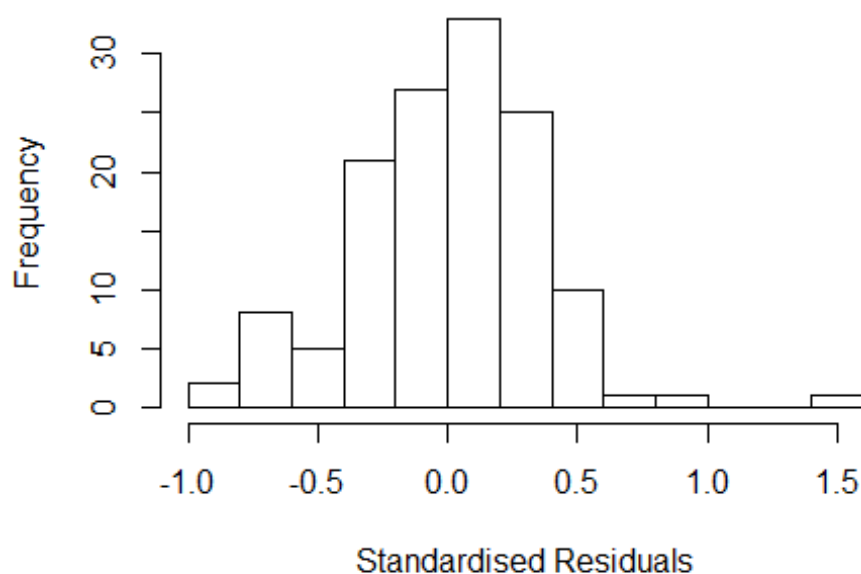
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9998 -0.1926 0.0185 0.1988 1.3202
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8916 0.5379 1.66 0.10
## FirstAuthorFemale1 -0.0396 0.0775 -0.51 0.61
## LastAuthorFemale1 -0.0454 0.0627 -0.72 0.47
## UniqueAuthors2 -0.2567 0.1866 -1.38 0.17
## UniqueAuthors3 -0.0111 0.1505 -0.07 0.94
## UniqueAuthors4 -0.1726 0.1333 -1.30 0.20
## UniqueAuthors5 -0.0113 0.1528 -0.07 0.94
## Year1997 0.4535 0.5562 0.82 0.42
## Year1998 0.6610 0.6231 1.06 0.29
## Year1999 0.2430 0.5913 0.41 0.68
```

```

## Year2000      0.1683      0.6249      0.27      0.79
## Year2001      0.4480      0.6152      0.73      0.47
## Year2002      0.5025      0.6465      0.78      0.44
## Year2003      0.2898      0.6331      0.46      0.65
## Year2004      0.2796      0.5850      0.48      0.63
## Year2005      0.1688      0.5877      0.29      0.77
## Year2006      0.1848      0.5745      0.32      0.75
## Year2007      0.3593      0.5805      0.62      0.54
## Year2008      0.2384      0.5587      0.43      0.67
## Year2009      0.4624      0.6137      0.75      0.45
## Year2010      0.1814      0.5866      0.31      0.76
## Year2011      0.1508      0.6238      0.24      0.81
## Year2012      0.3183      0.5782      0.55      0.58
##
## Robust residual standard error: 0.317
## Multiple R-squared:  0.23,   Adjusted R-squared:  0.0769
## Convergence in 36 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 116 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0449 0.8750 0.9560 0.8940 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      7.46e-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 2.691 1      1.640
## LastAuthorFemale 1.659 1      1.288
## Year      4.062 16      1.045

```

Residuals from first and last author



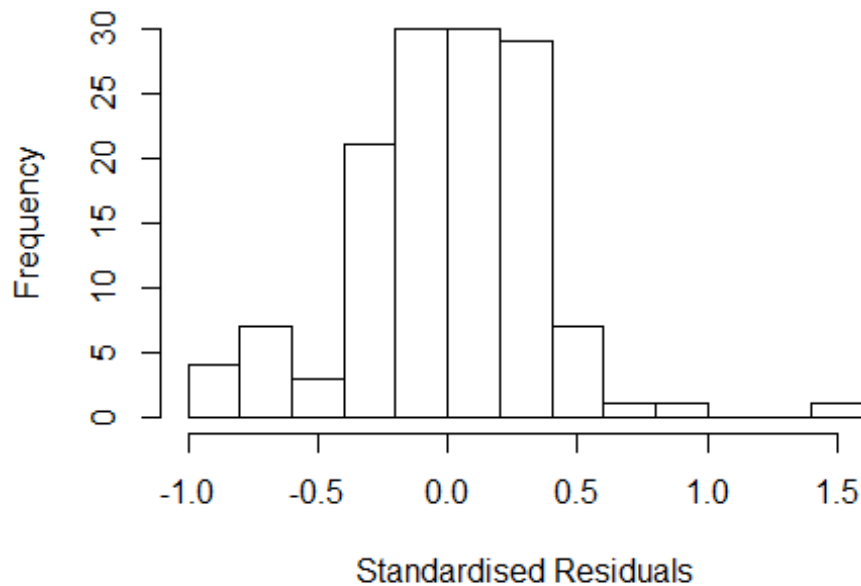
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9260 -0.2170 0.0261 0.2170 1.4638
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7226 0.7122 1.01 0.31
## FirstAuthorFemale1 -0.0254 0.0830 -0.31 0.76
## LastAuthorFemale1 -0.0708 0.0689 -1.03 0.31
## Year1997 0.5438 0.7112 0.76 0.45
## Year1998 0.8332 0.7389 1.13 0.26
## Year1999 0.3749 0.7176 0.52 0.60
## Year2000 0.2865 0.7116 0.40 0.69
## Year2001 0.4210 0.7442 0.57 0.57
## Year2002 0.5864 0.6836 0.86 0.39
## Year2003 0.4380 0.7148 0.61 0.54
## Year2004 0.3659 0.6982 0.52 0.60
## Year2005 0.2774 0.7115 0.39 0.70
```

```

## Year2006          0.2592      0.6954      0.37      0.71
## Year2007          0.4556      0.7071      0.64      0.52
## Year2008          0.3408      0.7083      0.48      0.63
## Year2009          0.5297      0.6996      0.76      0.45
## Year2010          0.2819      0.7021      0.40      0.69
## Year2011          0.1879      0.7088      0.27      0.79
## Year2012          0.4400      0.7158      0.61      0.54
##
## Robust residual standard error: 0.306
## Multiple R-squared:  0.182, Adjusted R-squared:  0.0542
## Convergence in 46 IRWLS iterations
##
## Robustness weights:
## observation 2 is an outlier with |weight| = 0 ( < 0.00075);
## 7 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.341  0.878  0.953  0.898  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      7.46e-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 2.231  1      1.494
## Year              2.231 16      1.025

```


Residuals from first author



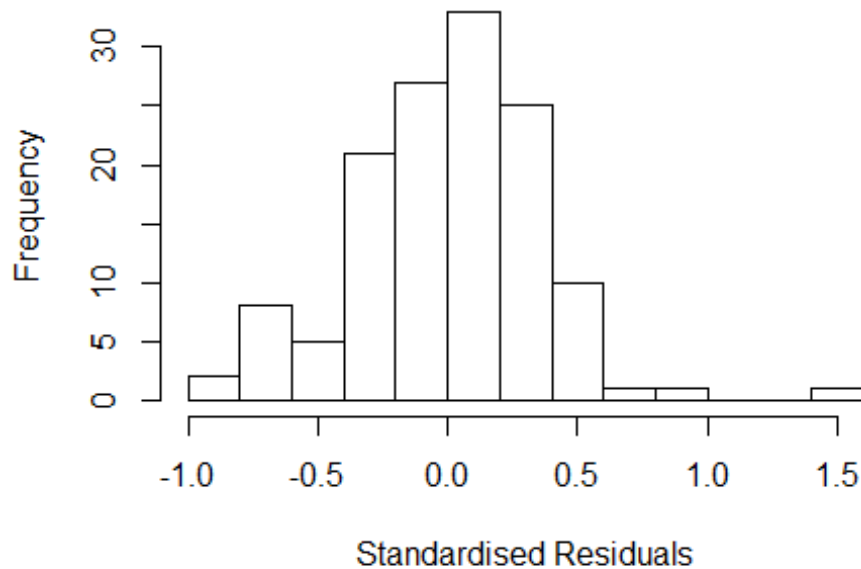
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9441 -0.2004 0.0204 0.2394 1.4912
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7034 0.6355 1.11 0.27
## FirstAuthorFemale1 -0.0337 0.0770 -0.44 0.66
## Year1997 0.5345 0.6307 0.85 0.40
## Year1998 0.8273 0.6856 1.21 0.23
## Year1999 0.3712 0.6418 0.58 0.56
## Year2000 0.2948 0.6451 0.46 0.65
## Year2001 0.4006 0.6686 0.60 0.55
## Year2002 0.5997 0.6370 0.94 0.35
## Year2003 0.4627 0.6484 0.71 0.48
## Year2004 0.3505 0.6235 0.56 0.58
## Year2005 0.3009 0.6402 0.47 0.64
## Year2006 0.2558 0.6251 0.41 0.68
```

```

## Year2007          0.4617      0.6298      0.73      0.46
## Year2008          0.3326      0.6236      0.53      0.59
## Year2009          0.5289      0.6281      0.84      0.40
## Year2010          0.2730      0.6234      0.44      0.66
## Year2011          0.2031      0.6307      0.32      0.75
## Year2012          0.4100      0.6329      0.65      0.52
##
## Robust residual standard error: 0.313
## Multiple R-squared: 0.172, Adjusted R-squared: 0.051
## Convergence in 38 IRWLS iterations
##
## Robustness weights:
## observation 2 is an outlier with |weight| = 0 ( < 0.00075);
## 9 weights are ~= 1. The remaining 124 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.342  0.890  0.942  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          7.46e-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 1.503 1          1.226
## Year             1.503 16          1.013

```

Residuals from last author



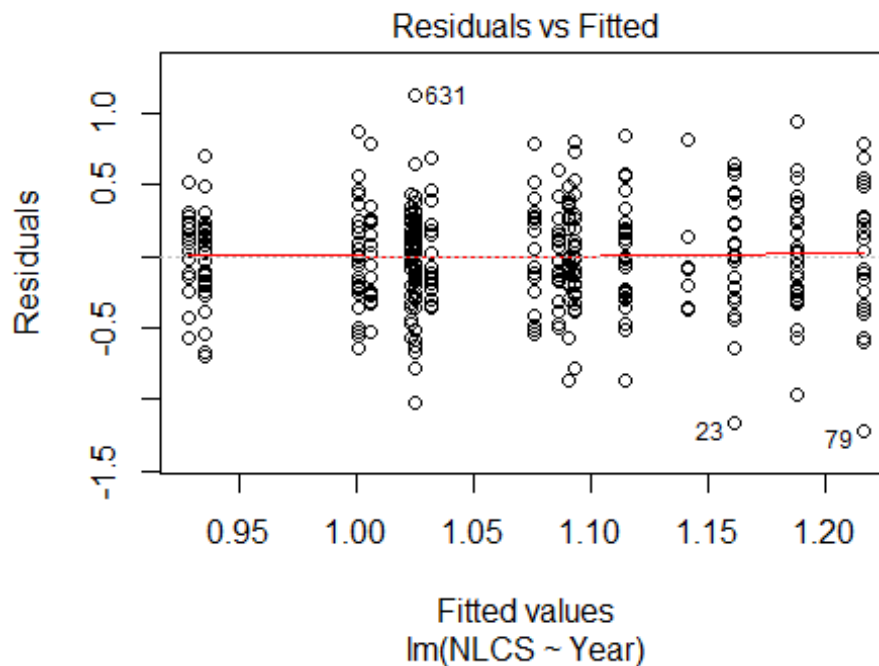
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.896 -0.215 0.033 0.224 1.464
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6968 0.6407 1.09 0.28
## LastAuthorFemale1 -0.0740 0.0667 -1.11 0.27
## Year1997 0.5653 0.6469 0.87 0.38
## Year1998 0.8521 0.6738 1.26 0.21
## Year1999 0.3982 0.6450 0.62 0.54
## Year2000 0.3035 0.6592 0.46 0.65
## Year2001 0.4426 0.6721 0.66 0.51
## Year2002 0.5817 0.6553 0.89 0.38
## Year2003 0.4459 0.6737 0.66 0.51
## Year2004 0.3771 0.6512 0.58 0.56
## Year2005 0.2923 0.6620 0.44 0.66
## Year2006 0.2694 0.6485 0.42 0.68
```

```

## Year2007          0.4747      0.6484      0.73      0.47
## Year2008          0.3587      0.6534      0.55      0.58
## Year2009          0.5404      0.6537      0.83      0.41
## Year2010          0.2977      0.6474      0.46      0.65
## Year2011          0.2003      0.6595      0.30      0.76
## Year2012          0.4602      0.6509      0.71      0.48
##
## Robust residual standard error: 0.317
## Multiple R-squared: 0.178, Adjusted R-squared: 0.0577
## Convergence in 42 IRWLS iterations
##
## Robustness weights:
## observation 2 is an outlier with |weight| <= 0.00074 ( < 0.00075);
## 6 weights are ~= 1. The remaining 127 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.405  0.888  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      7.46e-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: 134"
## [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 2008 2009 2010
##   36   48   34   35   42   51   43   29   38   39   49   44   44   44   44
## 2011 2012
##   47   41
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   24   21   21   19   17    8   27   16   26   24   29   30   20   30   36
## 2011 2012
##   34   30

```

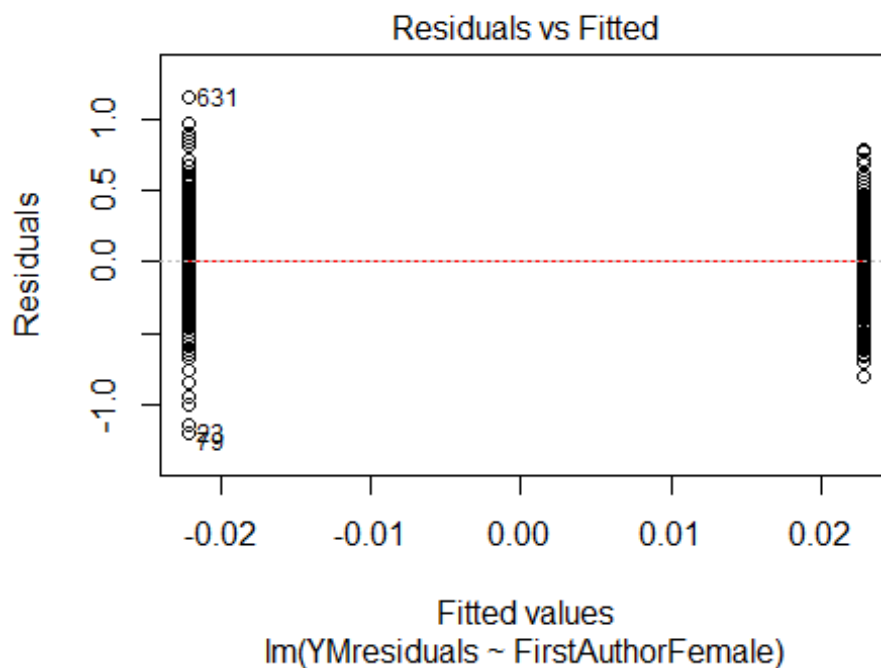
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   23   20   18   16   14    6   22   14   25   22   26   26   16   26   33
## 2011 2012
##   34   27
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity 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 = 9.1, df = 1, p-value = 0.003
## [1] "Female first author team size 2018 geometric mean: 4.12396580829147"
## [1] "Male first author team size 2018 geometric mean: 4.11209720697991"
## 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 = 60, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.04282321707986"
## [1] "Male last author team size 2018 geometric mean: 4.17821629545109"

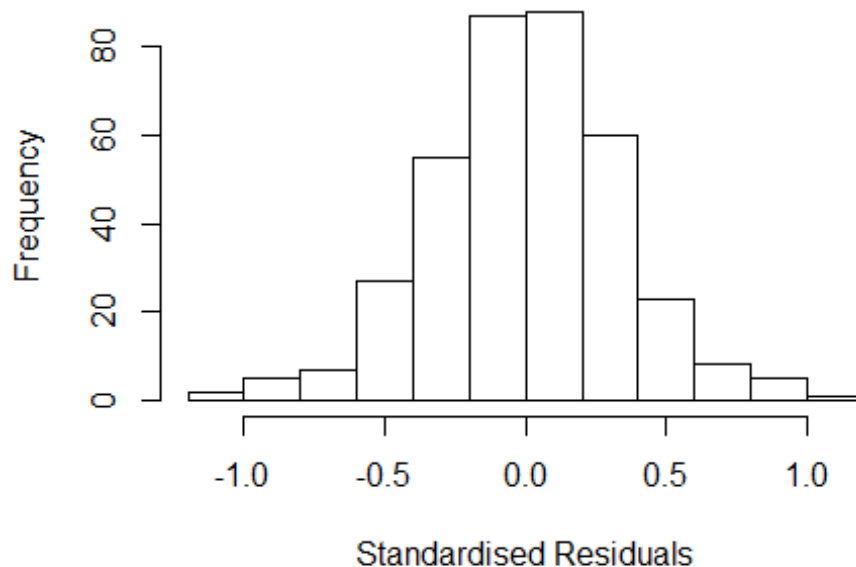
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 62, 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	1.406	1	1.186
LastAuthorFemale	1.427	1	1.195
UniqueAuthors	2.436	4	1.118
Year	3.421	16	1.039

Residuals from first and last author and team size



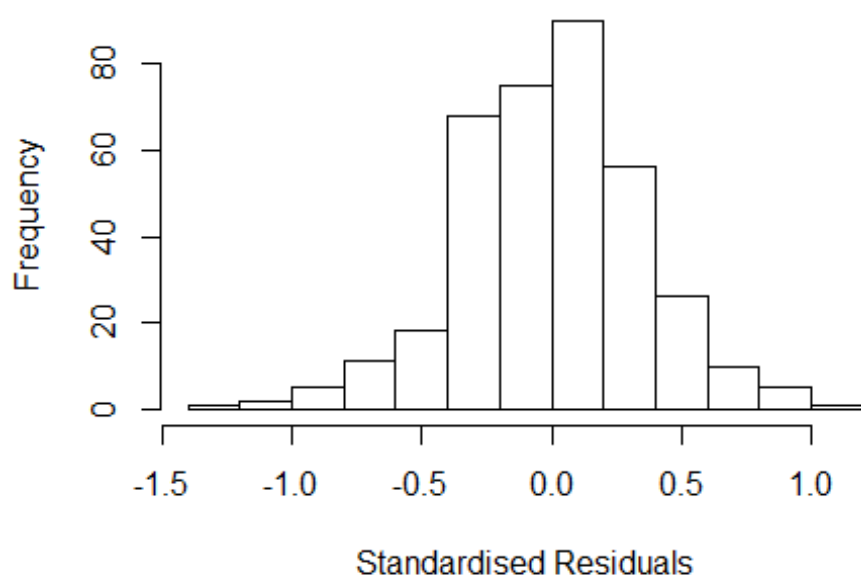
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.10361 -0.21171 0.00203 0.21509 1.04961
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9692 0.1417 6.84 3.7e-11 ***
## FirstAuthorFemale1 0.0490 0.0392 1.25 0.2121
## LastAuthorFemale1 0.0270 0.0414 0.65 0.5151
## UniqueAuthors2 0.1344 0.1357 0.99 0.3229
## UniqueAuthors3 0.1763 0.1358 1.30 0.1949
## UniqueAuthors4 0.2675 0.1345 1.99 0.0476 *
## UniqueAuthors5 0.3355 0.1336 2.51 0.0125 *
## Year1997 0.0369 0.1528 0.24 0.8093
## Year1998 -0.1710 0.1164 -1.47 0.1428
## Year1999 -0.2247 0.1278 -1.76 0.0797 .
```

```

## Year2000          -0.3081      0.1277    -2.41    0.0163 *
## Year2001          -0.1834      0.2018    -0.91    0.3639
## Year2002          -0.2671      0.1148    -2.33    0.0206 *
## Year2003          -0.2306      0.1258    -1.83    0.0675 .
## Year2004          -0.2060      0.1065    -1.93    0.0539 .
## Year2005          -0.1182      0.1120    -1.06    0.2921
## Year2006          -0.0460      0.1061    -0.43    0.6648
## Year2007          -0.2034      0.1247    -1.63    0.1038
## Year2008          -0.1841      0.1140    -1.62    0.1072
## Year2009          -0.2103      0.1028    -2.05    0.0416 *
## Year2010          -0.3292      0.1040    -3.16    0.0017 **
## Year2011          -0.1693      0.1087    -1.56    0.1201
## Year2012          -0.1138      0.1118    -1.02    0.3095
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.333
## Multiple R-squared:  0.116, Adjusted R-squared:  0.0597
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 339 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.248  0.886  0.957  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      2.72e-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 1.391 1      1.180
## LastAuthorFemale  1.273 1      1.128
## Year              1.619 16      1.015

```


Residuals from first and last author



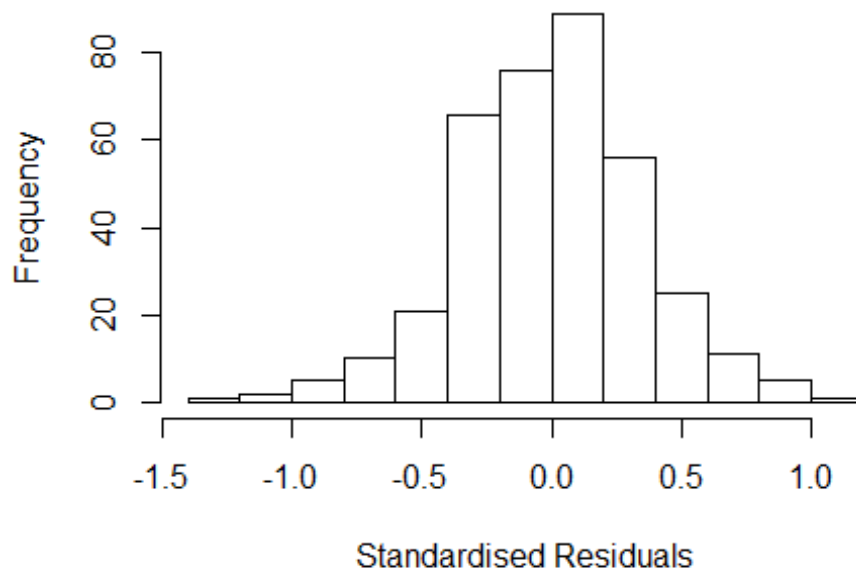
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20555 -0.24898 0.00787 0.21522 1.11448
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.14633 0.09548 12.01 <2e-16 ***
## FirstAuthorFemale1 0.05320 0.04035 1.32 0.188
## LastAuthorFemale1 0.01193 0.04088 0.29 0.771
## Year1997 0.05922 0.14462 0.41 0.682
## Year1998 -0.12195 0.11796 -1.03 0.302
## Year1999 -0.17724 0.13070 -1.36 0.176
## Year2000 -0.27478 0.12415 -2.21 0.028 *
## Year2001 -0.10781 0.17905 -0.60 0.547
## Year2002 -0.18774 0.12133 -1.55 0.123
## Year2003 -0.18952 0.12827 -1.48 0.140
## Year2004 -0.13214 0.10339 -1.28 0.202
## Year2005 -0.06260 0.10943 -0.57 0.568
```

```

## Year2006          0.00447    0.11146    0.04    0.968
## Year2007          -0.10981    0.12497   -0.88    0.380
## Year2008          -0.13529    0.11656   -1.16    0.247
## Year2009          -0.13541    0.10265   -1.32    0.188
## Year2010          -0.23526    0.09951   -2.36    0.019 *
## Year2011          -0.06712    0.10905   -0.62    0.539
## Year2012          -0.07989    0.11059   -0.72    0.471
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.0623, Adjusted R-squared:  0.014
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 344 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.168  0.883  0.952  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      2.72e-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 1.337 1      1.157
## Year              1.337 16      1.009

```

Residuals from first author



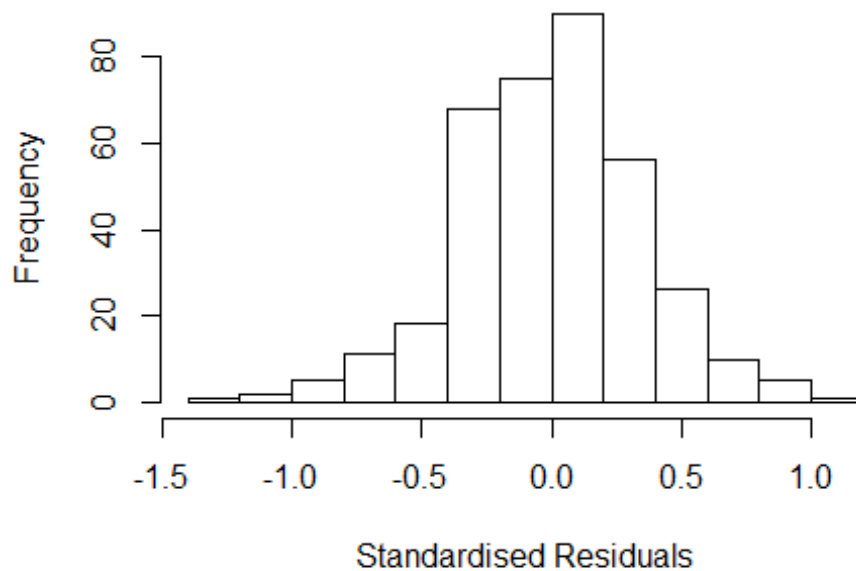
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2072 -0.2481 0.0101 0.2144 1.1093
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.15155 0.09240 12.46 <2e-16 ***
## FirstAuthorFemale1 0.05311 0.04025 1.32 0.188
## Year1997 0.05565 0.14266 0.39 0.697
## Year1998 -0.12554 0.11694 -1.07 0.284
## Year1999 -0.18170 0.12967 -1.40 0.162
## Year2000 -0.27750 0.12325 -2.25 0.025 *
## Year2001 -0.11198 0.17737 -0.63 0.528
## Year2002 -0.18955 0.12147 -1.56 0.120
## Year2003 -0.19009 0.12793 -1.49 0.138
## Year2004 -0.13289 0.10358 -1.28 0.200
## Year2005 -0.06469 0.10921 -0.59 0.554
## Year2006 0.00299 0.11169 0.03 0.979
```

```

## Year2007      -0.10982    0.12601   -0.87    0.384
## Year2008      -0.13827    0.11637   -1.19    0.236
## Year2009      -0.13582    0.10311   -1.32    0.189
## Year2010      -0.23668    0.09955   -2.38    0.018 *
## Year2011      -0.06513    0.10875   -0.60    0.550
## Year2012      -0.08118    0.11080   -0.73    0.464
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.332
## Multiple R-squared:  0.0625, Adjusted R-squared:  0.017
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 342 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.159  0.883  0.950  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      2.72e-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 1.219 1      1.104
## Year      1.219 16      1.006

```

Residuals from last author



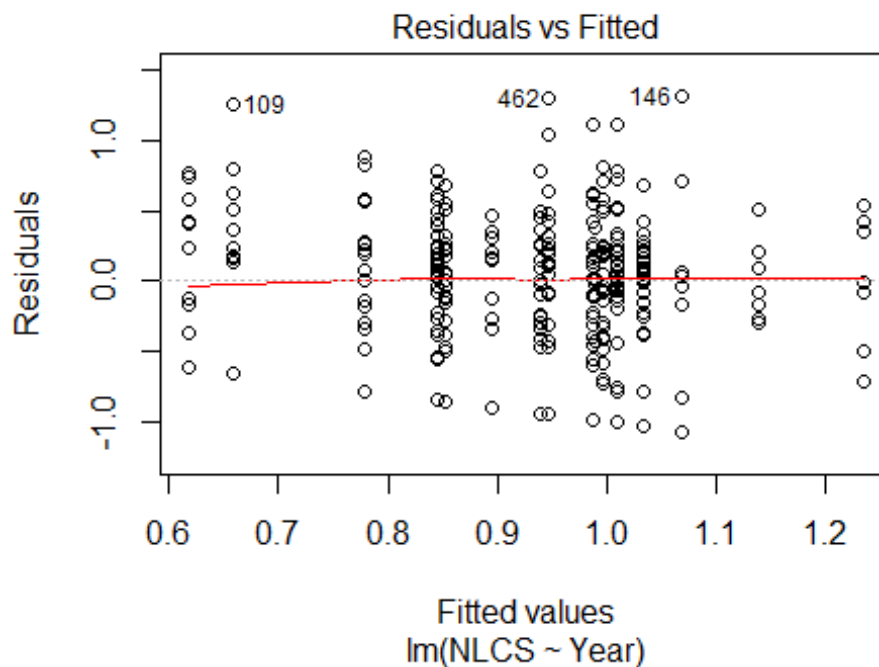
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.21911 -0.24479 -0.00191 0.22305 1.08871
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1701 0.0913 12.82 <2e-16 ***
## LastAuthorFemale1 0.0112 0.0406 0.28 0.783
## Year1997 0.0490 0.1434 0.34 0.733
## Year1998 -0.1178 0.1203 -0.98 0.328
## Year1999 -0.1856 0.1295 -1.43 0.153
## Year2000 -0.2731 0.1282 -2.13 0.034 *
## Year2001 -0.1170 0.1760 -0.66 0.507
## Year2002 -0.1849 0.1212 -1.53 0.128
## Year2003 -0.1717 0.1287 -1.33 0.183
## Year2004 -0.1377 0.1037 -1.33 0.185
## Year2005 -0.0564 0.1110 -0.51 0.612
## Year2006 0.0223 0.1149 0.19 0.846
```

```

## Year2007          -0.1078      0.1248   -0.86    0.388
## Year2008          -0.1328      0.1193   -1.11    0.266
## Year2009          -0.1347      0.1047   -1.29    0.199
## Year2010          -0.2308      0.1019   -2.26    0.024 *
## Year2011          -0.0483      0.1086   -0.45    0.656
## Year2012          -0.0784      0.1109   -0.71    0.480
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.0571, Adjusted R-squared:  0.0113
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 334 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.177  0.885   0.951   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      2.72e-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: 368"
## [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
##   13   12   13   24   38   30   18   23   28   23   39   30   31   38   31
## 2011 2012
##   31   57
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    7    7   11   17    8   10   13   21   21   30   26   22   26   24
## 2011 2012

```

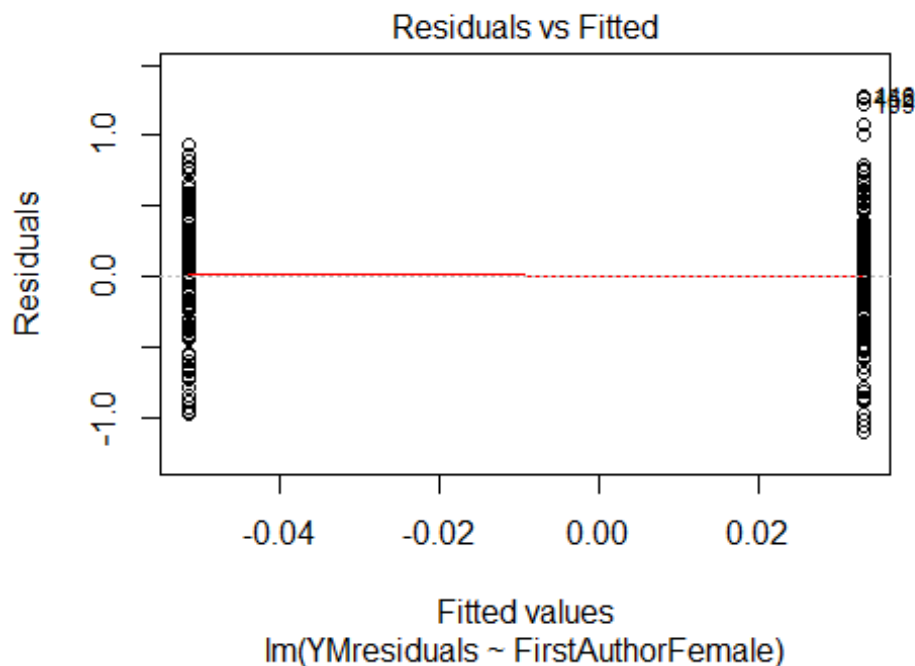
```
## 21 48
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 7 6 9 13 7 9 11 20 21 27 23 20 21 19
## 2011 2012
## 18 42
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 25, df = 16, p-value = 0.07
```



```
##
## 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: 4.28013680006884"
## [1] "Male first author team size 2018 geometric mean: 4.39429035136649"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

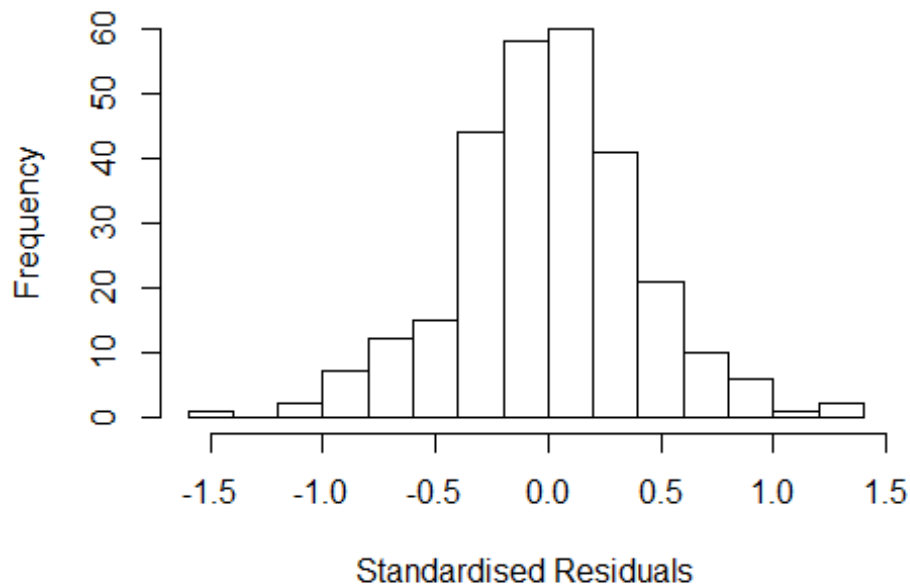
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 39, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.24534222319921"
## [1] "Male last author team size 2018 geometric mean: 4.81906978132569"

## 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 = 23, 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 1.395  1      1.181
## LastAuthorFemale  1.533  1      1.238
## UniqueAuthors    5.658  4      1.242
## Year              7.016 16      1.063
```


Residuals from first and last author and team size



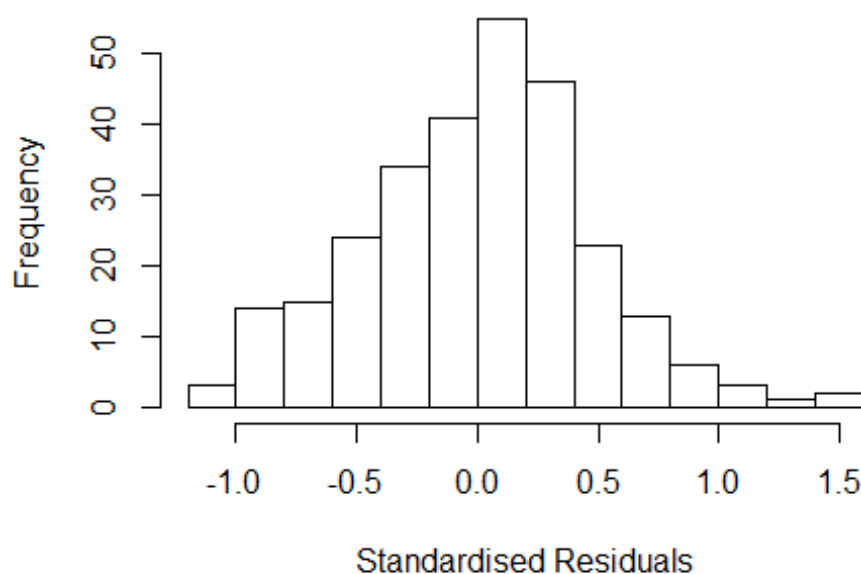
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.52260 -0.24129 0.00433 0.23191 1.24973
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4214 0.1293 3.26 0.0013 **
## FirstAuthorFemale1 -0.0859 0.0526 -1.63 0.1039
## LastAuthorFemale1 0.0689 0.0640 1.08 0.2828
## UniqueAuthors2 0.5917 0.1247 4.75 3.5e-06 ***
## UniqueAuthors3 0.5615 0.1002 5.60 5.4e-08 ***
## UniqueAuthors4 0.6502 0.1020 6.38 8.4e-10 ***
## UniqueAuthors5 0.8094 0.1026 7.89 8.6e-14 ***
## Year1997 0.1973 0.3254 0.61 0.5448
## Year1998 0.0942 0.1308 0.72 0.4724
## Year1999 -0.0305 0.1160 -0.26 0.7930
```

```

## Year2000          -0.1583      0.2602    -0.61    0.5435
## Year2001           0.5396      0.3420     1.58    0.1158
## Year2002          -0.0466      0.1244    -0.37    0.7081
## Year2003          -0.1480      0.1723    -0.86    0.3912
## Year2004          -0.1242      0.1117    -1.11    0.2675
## Year2005           0.0515      0.1073     0.48    0.6320
## Year2006          -0.0446      0.1235    -0.36    0.7185
## Year2007           0.0138      0.1261     0.11    0.9129
## Year2008          -0.0729      0.1109    -0.66    0.5118
## Year2009          -0.1180      0.1214    -0.97    0.3318
## Year2010          -0.0807      0.1459    -0.55    0.5805
## Year2011          -0.2848      0.1468    -1.94    0.0534 .
## Year2012          -0.1343      0.1093    -1.23    0.2203
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.355, Adjusted R-squared:  0.3
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 259 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0542 0.8680 0.9560 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          3.57e-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 1.268 1          1.126
## LastAuthorFemale 1.297 1          1.139
## Year          1.632 16          1.015

```

Residuals from first and last author



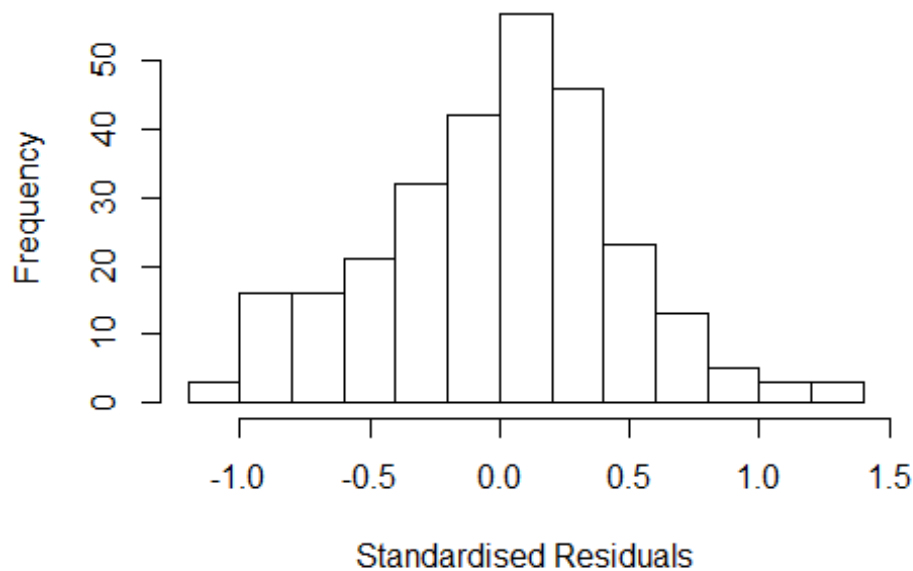
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1386 -0.3011 0.0297 0.3026 1.4341
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.028996 0.085336 12.06 <2e-16 ***
## FirstAuthorFemale1 -0.075821 0.059784 -1.27 0.206
## LastAuthorFemale1 -0.057267 0.076674 -0.75 0.456
## Year1997 0.242764 0.212184 1.14 0.254
## Year1998 0.058782 0.108635 0.54 0.589
## Year1999 0.000536 0.115475 0.00 0.996
## Year2000 -0.487852 0.197779 -2.47 0.014 *
## Year2001 -0.000989 0.454212 0.00 0.998
## Year2002 -0.069399 0.154689 -0.45 0.654
## Year2003 -0.310528 0.227469 -1.37 0.173
## Year2004 -0.118560 0.115560 -1.03 0.306
## Year2005 0.109631 0.123825 0.89 0.377
```

```

## Year2006      -0.007288    0.123327   -0.06    0.953
## Year2007      0.041535    0.122478    0.34    0.735
## Year2008     -0.067153    0.124712   -0.54    0.591
## Year2009     -0.055503    0.126542   -0.44    0.661
## Year2010     -0.086997    0.176433   -0.49    0.622
## Year2011     -0.232680    0.161380   -1.44    0.151
## Year2012     -0.108016    0.114342   -0.94    0.346
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.0965, Adjusted R-squared:  0.0342
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 252 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.272  0.869   0.952   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      3.57e-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 1.257 1      1.121
## Year              1.257 16      1.007

```

Residuals from first author



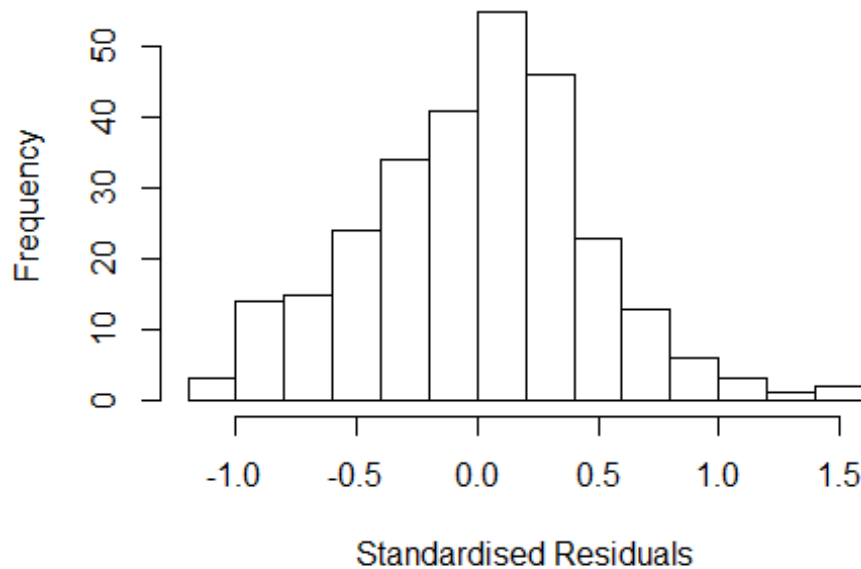
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1314 -0.3005 0.0362 0.2884 1.3880
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01510 0.08138 12.47 <2e-16 ***
## FirstAuthorFemale1 -0.08783 0.06199 -1.42 0.158
## Year1997 0.25319 0.20561 1.23 0.219
## Year1998 0.06693 0.10886 0.61 0.539
## Year1999 0.01571 0.11205 0.14 0.889
## Year2000 -0.48514 0.19708 -2.46 0.014 *
## Year2001 0.02817 0.44729 0.06 0.950
## Year2002 -0.07828 0.15563 -0.50 0.615
## Year2003 -0.30343 0.22997 -1.32 0.188
## Year2004 -0.11248 0.11564 -0.97 0.332
## Year2005 0.11630 0.12219 0.95 0.342
## Year2006 -0.00576 0.12289 -0.05 0.963
```

```

## Year2007          0.05357    0.12162    0.44    0.660
## Year2008          -0.05861    0.12234   -0.48    0.632
## Year2009          -0.04748    0.12528   -0.38    0.705
## Year2010          -0.09359    0.17314   -0.54    0.589
## Year2011          -0.23372    0.16292   -1.43    0.153
## Year2012          -0.10125    0.11428   -0.89    0.376
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.0941, Adjusted R-squared:  0.0353
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 255 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.302  0.870   0.949   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      3.57e-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 1.288 1      1.135
## Year              1.288 16      1.008

```

Residuals from last author



```
## [1] "List of 0 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
## -1.1026 -0.3264 0.0157 0.2892 1.4871
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01581 0.09072 11.20 <2e-16 ***
## LastAuthorFemale1 -0.08031 0.07920 -1.01 0.312
## Year1997 0.24557 0.21681 1.13 0.258
## Year1998 0.05009 0.11550 0.43 0.665
## Year1999 0.00599 0.12270 0.05 0.961
## Year2000 -0.50461 0.20784 -2.43 0.016 *
## Year2001 -0.00329 0.43452 -0.01 0.994
## Year2002 -0.09144 0.15446 -0.59 0.554
## Year2003 -0.31779 0.23753 -1.34 0.182
## Year2004 -0.12471 0.11999 -1.04 0.300
## Year2005 0.08679 0.12447 0.70 0.486
## Year2006 -0.02159 0.12679 -0.17 0.865
```

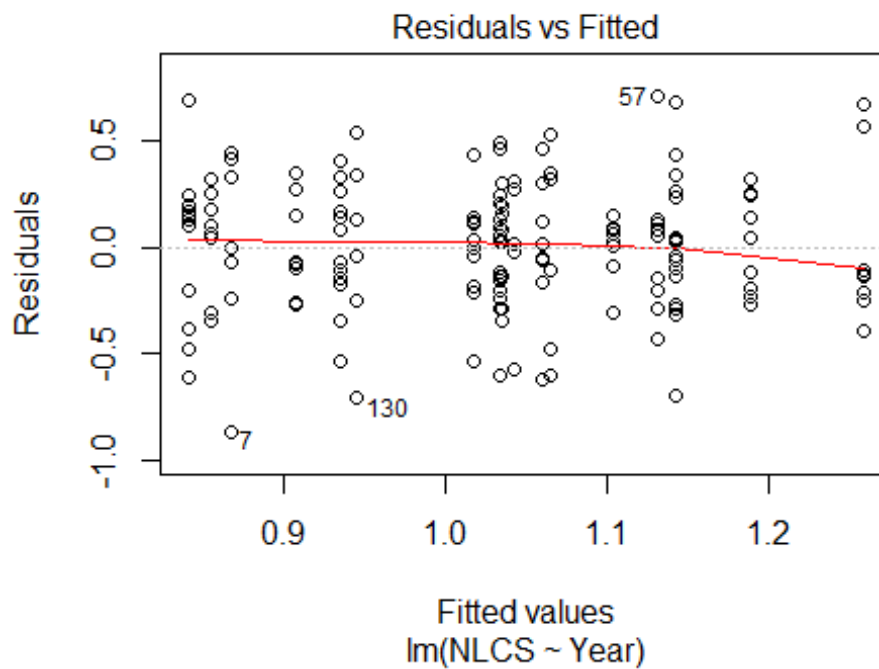
```

## Year2007      0.03605      0.12634      0.29      0.776
## Year2008     -0.07492      0.12685     -0.59      0.555
## Year2009     -0.07286      0.13310     -0.55      0.585
## Year2010     -0.09176      0.18214     -0.50      0.615
## Year2011     -0.24482      0.16288     -1.50      0.134
## Year2012     -0.12645      0.11496     -1.10      0.272
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.0929, Adjusted R-squared:  0.034
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 37 weights are ~= 1. The remaining 243 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.227  0.857  0.944  0.890  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      3.57e-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: 280"
## [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
##   21   21   23   13   26   14   17   15   15   12   21   20   10   17   18
## 2011 2012
##   25   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7   10    9    6    8    6   11    5    8    8   12   12    8   11   17
## 2011 2012

```



```
## 19 8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 8 5 5 7 5 11 2 7 8 10 11 7 10 16
## 2011 2012
## 18 8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



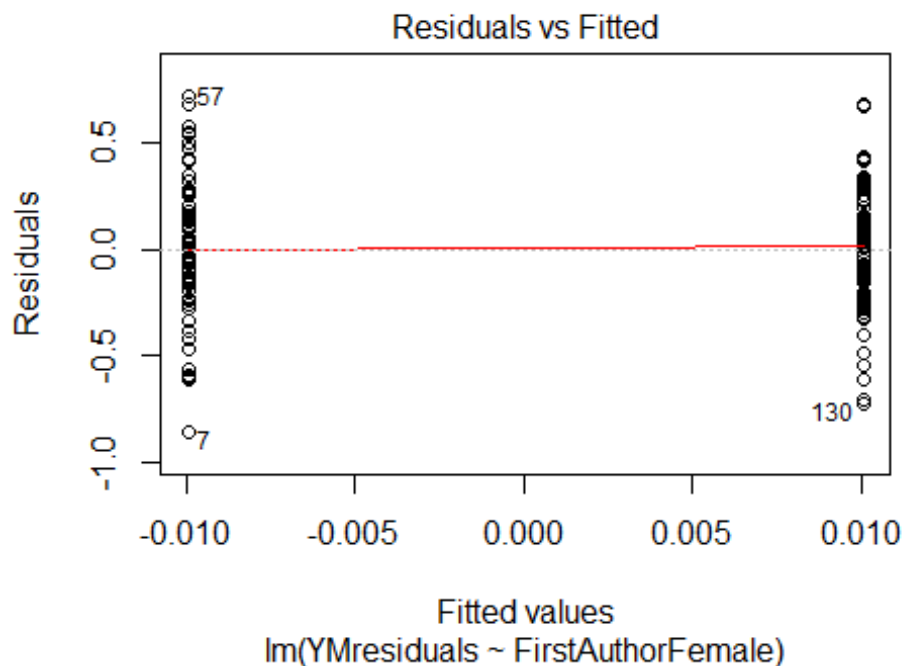
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.61, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 3.72791927319135"
## [1] "Male first author team size 2018 geometric mean: 3.46410161513775"

## 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 = 6, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.41995189335339"
## [1] "Male last author team size 2018 geometric mean: 3.83365862547764"

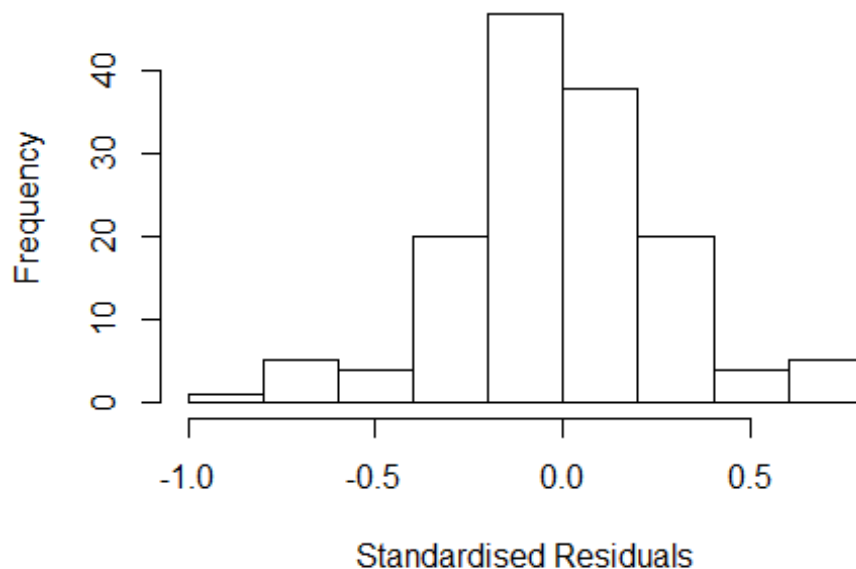
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, 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	3.350	1	1.830
LastAuthorFemale	1.663	1	1.290
UniqueAuthors	18.683	4	1.442
Year	42.599	16	1.124

Residuals from first and last author and team size



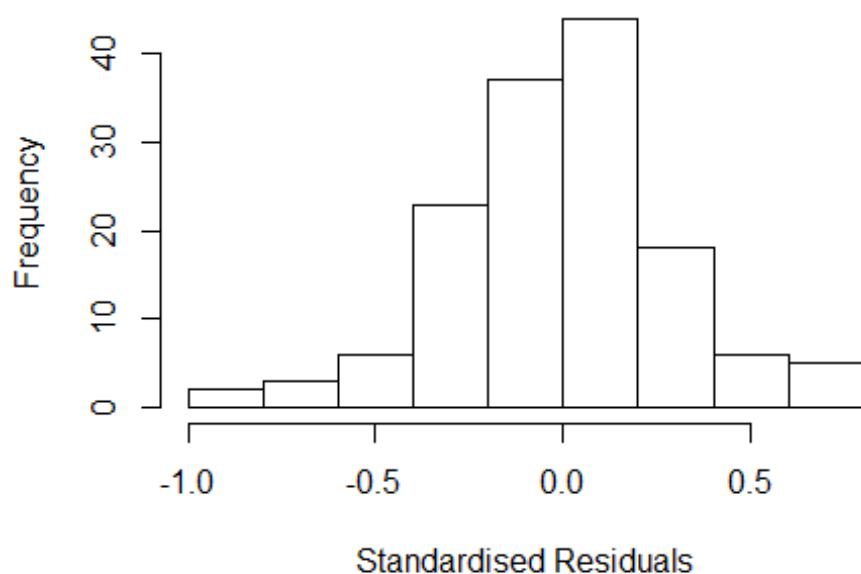
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9596 -0.1462 -0.0231 0.1657 0.7407
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.04493 0.19396 5.39 3.6e-07 ***
## FirstAuthorFemale1 0.03134 0.06839 0.46 0.65
## LastAuthorFemale1 -0.05343 0.05847 -0.91 0.36
## UniqueAuthors2 -0.03186 0.09583 -0.33 0.74
## UniqueAuthors3 0.08153 0.09175 0.89 0.38
## UniqueAuthors4 -0.00647 0.09686 -0.07 0.95
## UniqueAuthors5 0.11304 0.10520 1.07 0.28
## Year1997 -0.25297 0.23132 -1.09 0.28
## Year1998 -0.11394 0.22986 -0.50 0.62
## Year1999 -0.16448 0.60578 -0.27 0.79
```

```

## Year2000      -0.08661    0.22760   -0.38    0.70
## Year2001      0.03581    0.30628    0.12    0.91
## Year2002     -0.19617    0.25305   -0.78    0.44
## Year2003     -0.06012    0.22159   -0.27    0.79
## Year2004      0.03603    0.21564    0.17    0.87
## Year2005      0.00867    0.24632    0.04    0.97
## Year2006      0.10399    0.21817    0.48    0.63
## Year2007     -0.22863    0.22760   -1.00    0.32
## Year2008     -0.15606    0.22464   -0.69    0.49
## Year2009     -0.07723    0.21672   -0.36    0.72
## Year2010     -0.05238    0.21827   -0.24    0.81
## Year2011      0.02821    0.22437    0.13    0.90
## Year2012      0.09927    0.27904    0.36    0.72
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.258
## Multiple R-squared:  0.19,   Adjusted R-squared:  0.0432
## Convergence in 33 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 138 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.138  0.876  0.961  0.896  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      6.94e-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 2.385 1      1.544
## LastAuthorFemale  1.527 1      1.236
## Year              3.421 16      1.039

```

Residuals from first and last author



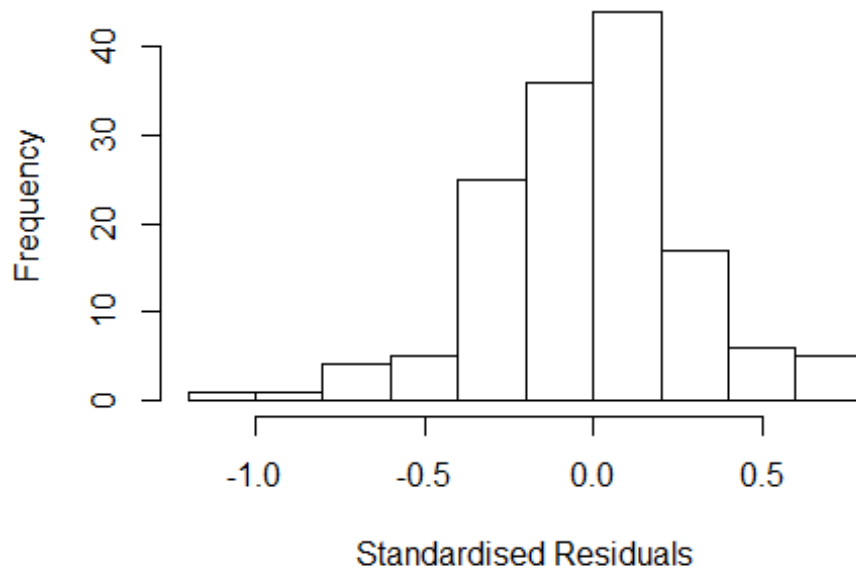
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.97256 -0.17169 0.00946 0.15428 0.78460
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03463 0.19642 5.27 5.9e-07 ***
## FirstAuthorFemale1 0.04444 0.06197 0.72 0.47
## LastAuthorFemale1 -0.06207 0.05512 -1.13 0.26
## Year1997 -0.22959 0.20972 -1.09 0.28
## Year1998 -0.10428 0.20418 -0.51 0.61
## Year1999 -0.06881 0.49743 -0.14 0.89
## Year2000 -0.03245 0.20142 -0.16 0.87
## Year2001 0.13818 0.25102 0.55 0.58
## Year2002 -0.16152 0.24861 -0.65 0.52
## Year2003 -0.01886 0.19039 -0.10 0.92
## Year2004 0.05864 0.19319 0.30 0.76
## Year2005 0.05124 0.21669 0.24 0.81
```

```

## Year2006          0.14277    0.19593    0.73    0.47
## Year2007          -0.14789    0.20133   -0.73    0.46
## Year2008          -0.12822    0.22325   -0.57    0.57
## Year2009          -0.02948    0.19695   -0.15    0.88
## Year2010           0.00767    0.19762    0.04    0.97
## Year2011           0.08537    0.20217    0.42    0.67
## Year2012           0.10976    0.23333    0.47    0.64
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.255
## Multiple R-squared:  0.156, Adjusted R-squared:  0.0342
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 133 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.112  0.885   0.955   0.885   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      6.94e-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 2.21 1      1.487
## Year              2.21 16      1.025

```

Residuals from first author



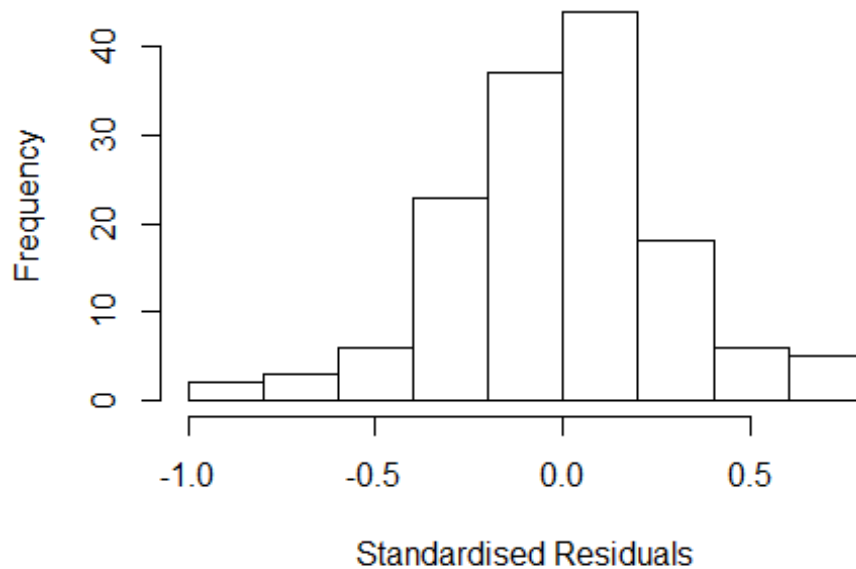
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.00158 -0.19413 0.00111 0.16440 0.78848
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0016 0.2216 4.52 1.4e-05 ***
## FirstAuthorFemale1 0.0350 0.0609 0.57 0.57
## Year1997 -0.2138 0.2305 -0.93 0.36
## Year1998 -0.0831 0.2292 -0.36 0.72
## Year1999 -0.0393 0.4489 -0.09 0.93
## Year2000 -0.0126 0.2271 -0.06 0.96
## Year2001 0.1323 0.2880 0.46 0.65
## Year2002 -0.1608 0.2523 -0.64 0.53
## Year2003 0.0189 0.2144 0.09 0.93
## Year2004 0.0787 0.2166 0.36 0.72
## Year2005 0.0755 0.2409 0.31 0.75
## Year2006 0.1672 0.2214 0.76 0.45
```

```

## Year2007          -0.1154      0.2240    -0.52      0.61
## Year2008          -0.1067      0.2340    -0.46      0.65
## Year2009          -0.0184      0.2218    -0.08      0.93
## Year2010           0.0326      0.2208      0.15      0.88
## Year2011           0.0982      0.2251      0.44      0.66
## Year2012           0.1389      0.2560      0.54      0.59
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.263
## Multiple R-squared:  0.146, Adjusted R-squared:  0.0304
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 131 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.117  0.885   0.953   0.889   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      6.94e-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 1.739 1          1.319
## Year              1.739 16          1.017

```


Residuals from last author



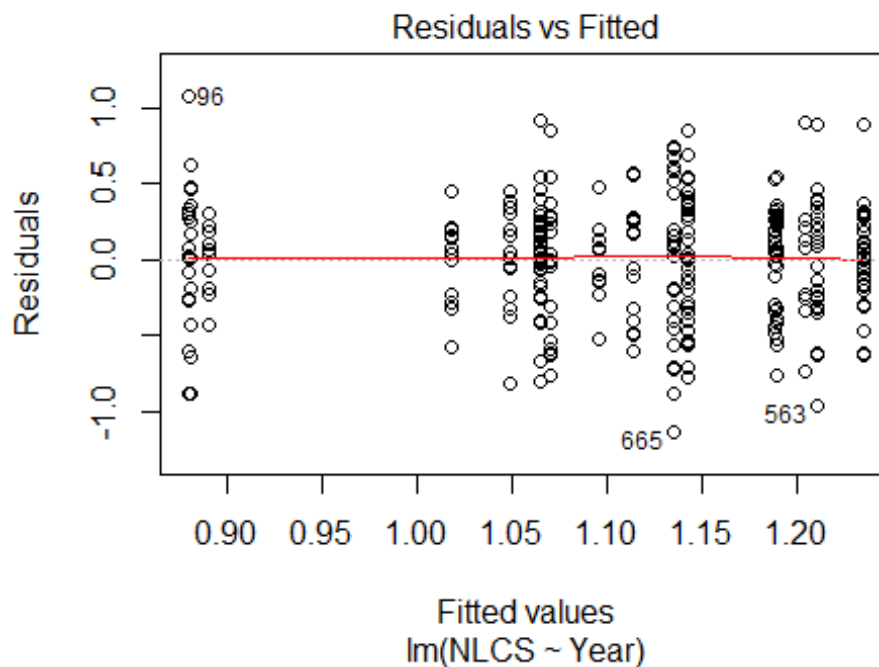
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.008 -0.180 0.009 0.152 0.752
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06556 0.18721 5.69 8.4e-08 ***
## LastAuthorFemale1 -0.05725 0.05519 -1.04 0.30
## Year1997 -0.25084 0.20437 -1.23 0.22
## Year1998 -0.12847 0.20017 -0.64 0.52
## Year1999 -0.09256 0.51524 -0.18 0.86
## Year2000 -0.03985 0.20319 -0.20 0.84
## Year2001 0.08817 0.23610 0.37 0.71
## Year2002 -0.18374 0.23040 -0.80 0.43
## Year2003 -0.02756 0.18760 -0.15 0.88
## Year2004 0.05726 0.19282 0.30 0.77
## Year2005 0.02509 0.21307 0.12 0.91
## Year2006 0.13284 0.19589 0.68 0.50
```

```

## Year2007          -0.15378      0.20315    -0.76      0.45
## Year2008          -0.13803      0.22169    -0.62      0.53
## Year2009          -0.02385      0.19882    -0.12      0.90
## Year2010           0.00461      0.19794      0.02      0.98
## Year2011           0.07277      0.19964      0.36      0.72
## Year2012           0.11099      0.23866      0.47      0.64
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.258
## Multiple R-squared:  0.151, Adjusted R-squared:  0.0363
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 131 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0911 0.8820 0.9510 0.8860 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      6.94e-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: 144"
## [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
##   40   48   47   39   34   41   37   26   26   32   43   42   55   36   34
## 2011 2012
##   42   37
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   15   12   11   11    9   11   16   12   13   18   31   26   33   25   21
## 2011 2012

```

```
## 22 21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 15 12 10 10 5 11 16 7 12 14 26 23 30 23 18
## 2011 2012
## 21 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 = 26, df = 16, p-value = 0.05
```



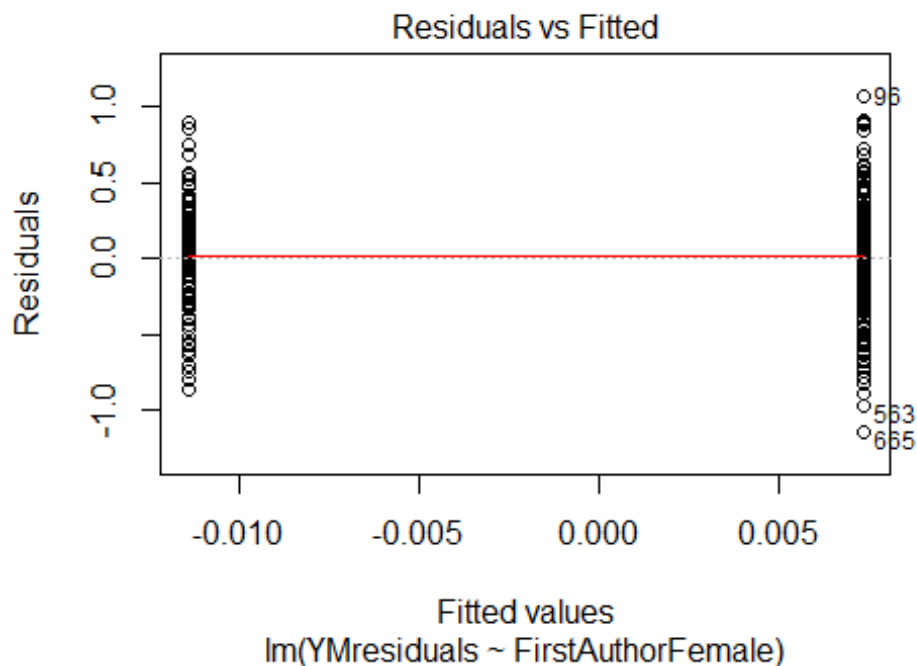
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.052, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 3.94455744367802"
## [1] "Male first author team size 2018 geometric mean: 1.90365393871588"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

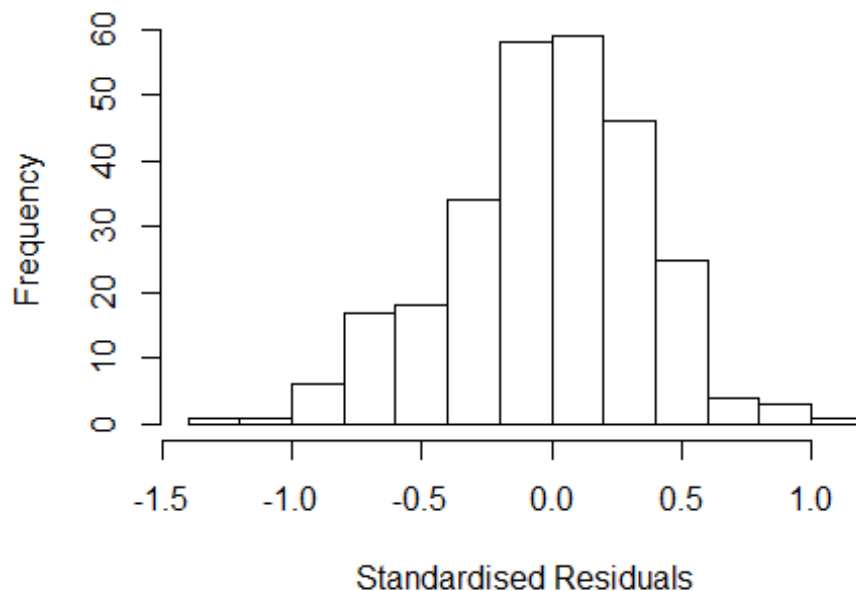
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 54, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.78176249895018"
## [1] "Male last author team size 2018 geometric mean: 2.90178313071286"

## 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 = 0.1
## 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 1.319  1      1.148
## LastAuthorFemale  1.806  1      1.344
## UniqueAuthors    5.138  4      1.227
## Year              9.340 16      1.072
```

Residuals from first and last author and team size



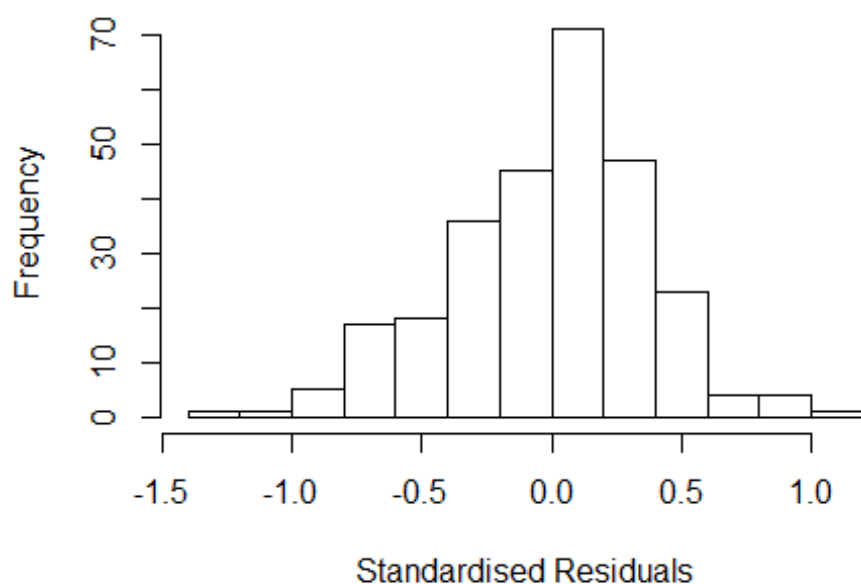
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2760 -0.2452 0.0156 0.2438 1.1049
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9155 0.1316 6.96 3e-11 ***
## FirstAuthorFemale1 -0.0319 0.0503 -0.63 0.526
## LastAuthorFemale1 0.0445 0.0628 0.71 0.480
## UniqueAuthors2 0.1891 0.1514 1.25 0.213
## UniqueAuthors3 0.2097 0.1287 1.63 0.104
## UniqueAuthors4 0.2384 0.1242 1.92 0.056 .
## UniqueAuthors5 0.2803 0.1176 2.38 0.018 *
## Year1997 -0.2970 0.1661 -1.79 0.075 .
## Year1998 -0.0527 0.1756 -0.30 0.764
## Year1999 -0.2150 0.1154 -1.86 0.064 .
```

```

## Year2000      0.1502      0.2443      0.61      0.539
## Year2001     -0.0819      0.1390     -0.59      0.556
## Year2002     -0.0956      0.1344     -0.71      0.478
## Year2003     -0.2500      0.2784     -0.90      0.370
## Year2004     -0.1047      0.1328     -0.79      0.431
## Year2005      0.1245      0.1283      0.97      0.333
## Year2006      0.0787      0.1214      0.65      0.517
## Year2007      0.0771      0.1326      0.58      0.562
## Year2008     -0.0902      0.1186     -0.76      0.448
## Year2009      0.0801      0.1727      0.46      0.643
## Year2010      0.0475      0.1579      0.30      0.764
## Year2011     -0.0484      0.1505     -0.32      0.748
## Year2012      0.0622      0.1416      0.44      0.661
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0338
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 246 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.158  0.856  0.947  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      3.66e-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 1.286 1      1.134
## LastAuthorFemale  1.477 1      1.215
## Year              1.865 16      1.020

```

Residuals from first and last author



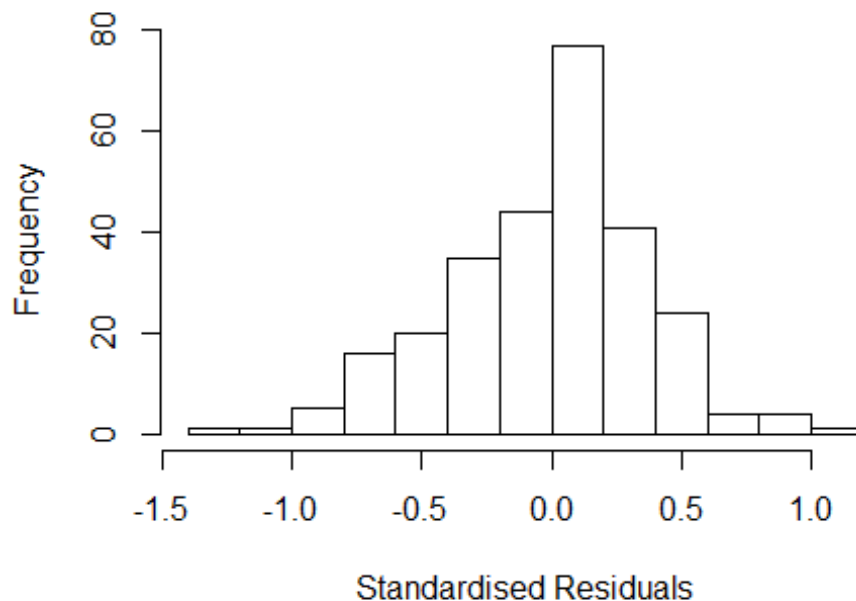
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2330 -0.2553  0.0279  0.2293  1.0602
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.12047    0.11364   9.86  <2e-16 ***
## FirstAuthorFemale1 -0.02805    0.05061  -0.55    0.58
## LastAuthorFemale1  0.03349    0.06058   0.55    0.58
## Year1997        -0.25719    0.17338  -1.48    0.14
## Year1998        -0.02542    0.18789  -0.14    0.89
## Year1999        -0.19736    0.13045  -1.51    0.13
## Year2000         0.02017    0.21355   0.09    0.92
## Year2001        -0.03298    0.14254  -0.23    0.82
## Year2002        -0.05642    0.13924  -0.41    0.69
## Year2003        -0.20105    0.29507  -0.68    0.50
## Year2004        -0.06782    0.14067  -0.48    0.63
## Year2005         0.15464    0.13617   1.14    0.26
```

```

## Year2006          0.12842    0.12828    1.00    0.32
## Year2007          0.12011    0.14208    0.85    0.40
## Year2008         -0.06990    0.12983   -0.54    0.59
## Year2009          0.11252    0.18911    0.59    0.55
## Year2010          0.08868    0.16309    0.54    0.59
## Year2011          0.00893    0.15451    0.06    0.95
## Year2012          0.09923    0.14389    0.69    0.49
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.353
## Multiple R-squared:  0.0839, Adjusted R-squared:  0.019
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 245 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.198  0.858  0.948  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      3.66e-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 1.273 1         1.128
## Year              1.273 16         1.008

```


Residuals from first author



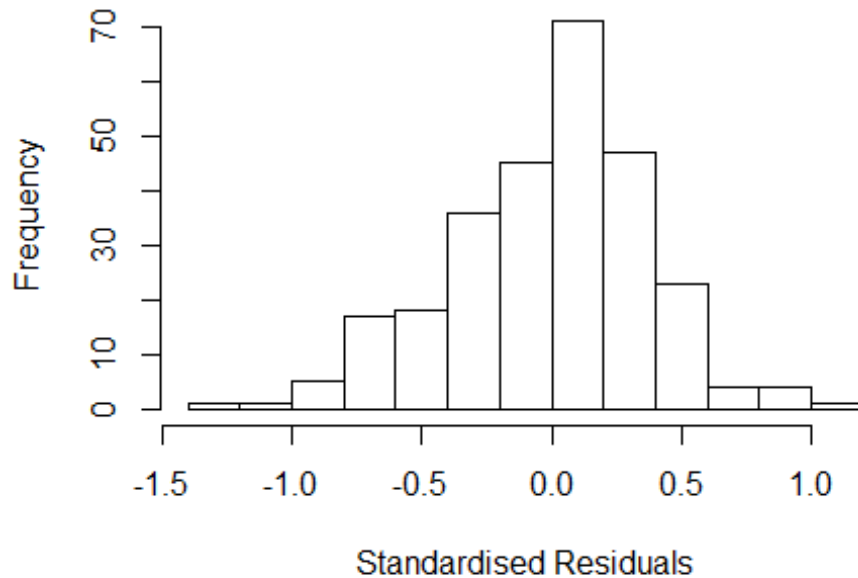
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2363 -0.2469 0.0225 0.2253 1.0967
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1242 0.1113 10.10 <2e-16 ***
## FirstAuthorFemale1 -0.0270 0.0505 -0.53 0.59
## Year1997 -0.2639 0.1705 -1.55 0.12
## Year1998 -0.0269 0.1875 -0.14 0.89
## Year1999 -0.2013 0.1284 -1.57 0.12
## Year2000 0.0166 0.2124 0.08 0.94
## Year2001 -0.0205 0.1365 -0.15 0.88
## Year2002 -0.0458 0.1363 -0.34 0.74
## Year2003 -0.1848 0.2877 -0.64 0.52
## Year2004 -0.0605 0.1386 -0.44 0.66
## Year2005 0.1583 0.1331 1.19 0.24
## Year2006 0.1304 0.1256 1.04 0.30
```

```

## Year2007          0.1227      0.1398      0.88      0.38
## Year2008         -0.0664      0.1267     -0.52      0.60
## Year2009          0.1121      0.1875      0.60      0.55
## Year2010          0.0960      0.1592      0.60      0.55
## Year2011          0.0138      0.1527      0.09      0.93
## Year2012          0.0986      0.1419      0.69      0.49
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.354
## Multiple R-squared:  0.0828, Adjusted R-squared:  0.0216
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 247 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.198  0.858  0.948  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      3.66e-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 1.465 1          1.210
## Year            1.465 16          1.012

```

Residuals from last author



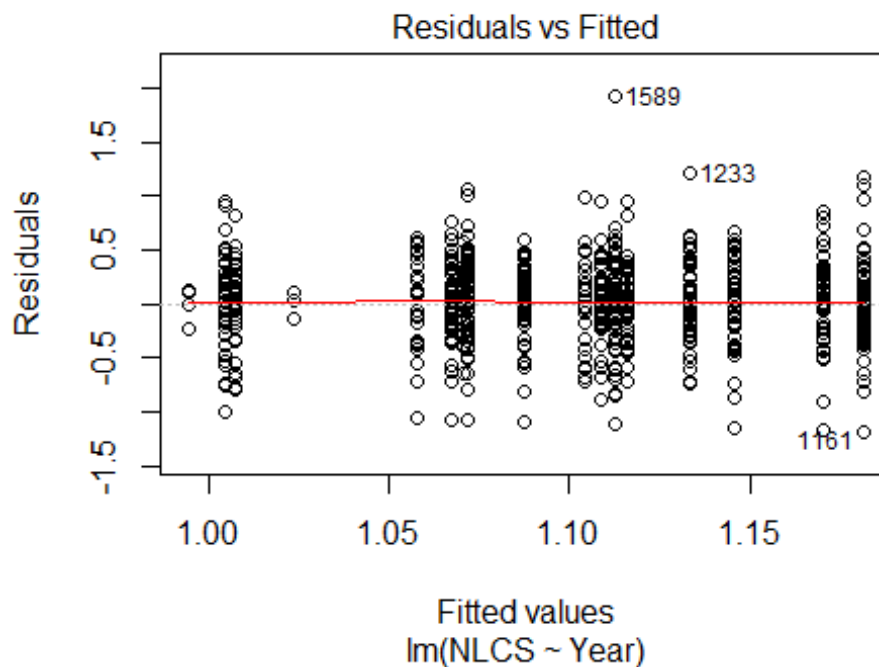
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2234 -0.2544 0.0316 0.2288 1.0656
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.11777 0.11389 9.81 <2e-16 ***
## LastAuthorFemale1 0.03208 0.06023 0.53 0.59
## Year1997 -0.25844 0.17347 -1.49 0.14
## Year1998 -0.03100 0.18587 -0.17 0.87
## Year1999 -0.19769 0.13048 -1.52 0.13
## Year2000 0.00595 0.21205 0.03 0.98
## Year2001 -0.03533 0.14198 -0.25 0.80
## Year2002 -0.06220 0.13854 -0.45 0.65
## Year2003 -0.21183 0.30124 -0.70 0.48
## Year2004 -0.07880 0.13844 -0.57 0.57
## Year2005 0.14440 0.13317 1.08 0.28
## Year2006 0.12150 0.12774 0.95 0.34
```

```

## Year2007      0.10867      0.13843      0.79      0.43
## Year2008     -0.07994      0.12847     -0.62      0.53
## Year2009      0.10564      0.19103      0.55      0.58
## Year2010      0.08086      0.16271      0.50      0.62
## Year2011     -0.00805      0.15289     -0.05      0.96
## Year2012      0.08840      0.14246      0.62      0.54
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.0833, Adjusted R-squared:  0.0222
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 244 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.202  0.857  0.946  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      3.66e-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: 273"
## [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
## 114 118 8 84 12 72 84 61 59 70 89 101 105 87 108
## 2011 2012
## 82 96
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 46 51 3 36 4 26 48 31 31 43 57 58 60 50 67
## 2011 2012

```

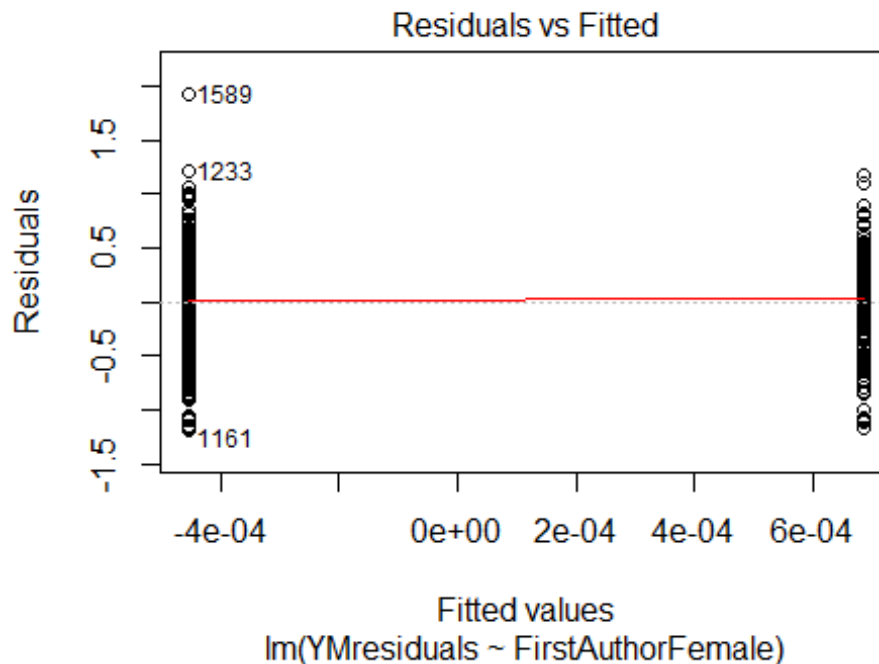
```
## 47 58
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 43 47 3 31 4 20 42 29 28 39 52 51 55 46 64
## 2011 2012
## 45 51
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 2.3, df = 1, p-value = 0.1
## [1] "Female first author team size 2018 geometric mean: 3.75580403106665"
## [1] "Male first author team size 2018 geometric mean: 5.43828473728978"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

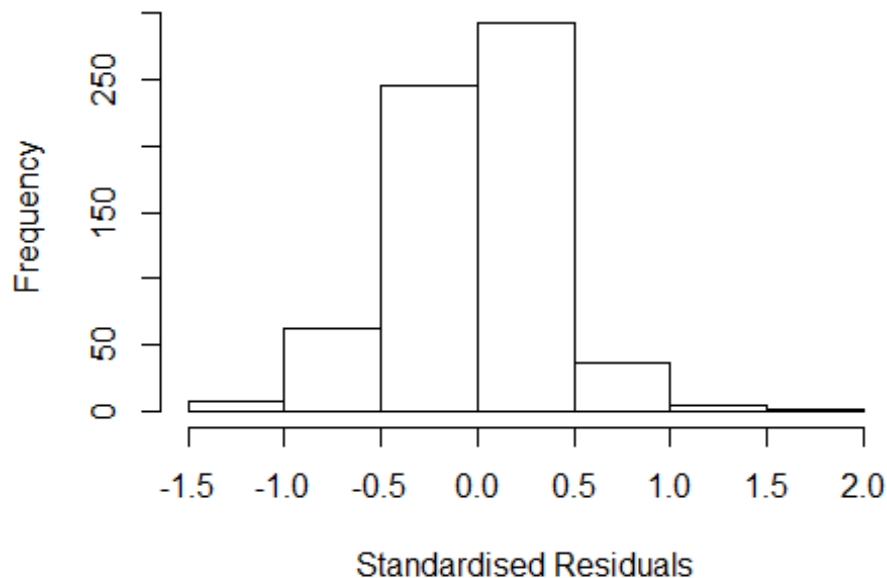
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120, p-value = 0.05
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.84577637271709"
## [1] "Male last author team size 2018 geometric mean: 5.0564432988773"

## 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 = 120, 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 1.375  1      1.173
## LastAuthorFemale  1.172  1      1.083
## UniqueAuthors    2.121  4      1.099
## Year              2.583 16      1.030
```

Residuals from first and last author and team size



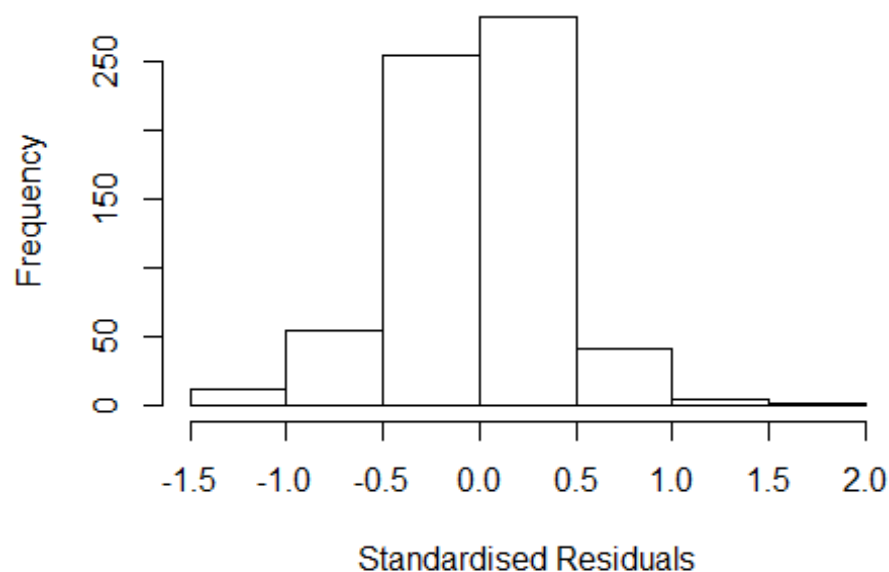
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.18253 -0.22278 0.00853 0.23395 1.96140
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.05e+00 7.32e-02 14.38 < 2e-16 ***
## FirstAuthorFemale1 6.53e-05 3.29e-02 0.00 0.99842
## LastAuthorFemale1 7.08e-03 3.65e-02 0.19 0.84619
## UniqueAuthors2 1.23e-01 5.63e-02 2.18 0.02956 *
## UniqueAuthors3 1.62e-01 5.70e-02 2.85 0.00457 **
## UniqueAuthors4 1.60e-01 5.76e-02 2.78 0.00559 **
## UniqueAuthors5 2.07e-01 5.66e-02 3.66 0.00027 ***
## Year1997 7.40e-03 8.48e-02 0.09 0.93050
## Year1998 -8.51e-02 8.99e-02 -0.95 0.34459
## Year1999 -1.49e-01 9.16e-02 -1.63 0.10401
```

```

## Year2000      -2.22e-01  9.68e-02  -2.30  0.02197 *
## Year2001      -1.37e-01  1.50e-01  -0.91  0.36058
## Year2002      -1.29e-01  7.78e-02  -1.66  0.09687 .
## Year2003      -1.18e-01  7.81e-02  -1.52  0.12981
## Year2004      -5.75e-02  9.58e-02  -0.60  0.54861
## Year2005      -6.36e-02  8.15e-02  -0.78  0.43559
## Year2006      -9.60e-02  7.94e-02  -1.21  0.22736
## Year2007      -3.37e-02  8.07e-02  -0.42  0.67667
## Year2008      -7.97e-02  7.91e-02  -1.01  0.31435
## Year2009      -9.14e-02  8.29e-02  -1.10  0.27070
## Year2010      -1.03e-01  8.85e-02  -1.17  0.24258
## Year2011      -1.72e-01  8.93e-02  -1.93  0.05462 .
## Year2012      -1.14e-01  8.73e-02  -1.31  0.19075
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.348
## Multiple R-squared:  0.0413, Adjusted R-squared:  0.00766
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 541 is an outlier with |weight| = 0 ( < 0.00015);
## 61 weights are ~= 1. The remaining 588 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.176  0.855  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           1.54e-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 1.232 1 1.110
## LastAuthorFemale 1.130 1 1.063
## Year 1.369 16 1.010

```


Residuals from first and last author



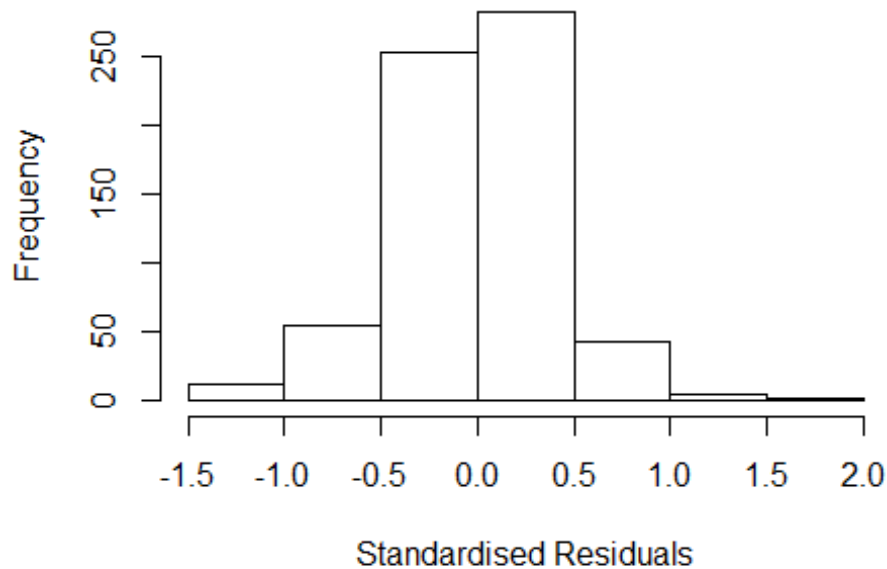
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.21718 -0.22680 0.00994 0.24143 1.93504
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1767 0.0678 17.36 <2e-16 ***
## FirstAuthorFemale1 0.0228 0.0317 0.72 0.473
## LastAuthorFemale1 0.0144 0.0366 0.39 0.693
## Year1997 0.0177 0.0879 0.20 0.841
## Year1998 -0.1728 0.0810 -2.13 0.033 *
## Year1999 -0.1225 0.0934 -1.31 0.190
## Year2000 -0.2076 0.0957 -2.17 0.030 *
## Year2001 -0.1243 0.1497 -0.83 0.406
## Year2002 -0.1261 0.0807 -1.56 0.119
## Year2003 -0.1168 0.0784 -1.49 0.137
## Year2004 -0.0563 0.0963 -0.58 0.559
## Year2005 -0.0507 0.0824 -0.62 0.539
```

```

## Year2006          -0.0863      0.0806   -1.07    0.284
## Year2007          -0.0245      0.0829   -0.30    0.767
## Year2008          -0.0751      0.0811   -0.93    0.355
## Year2009          -0.0692      0.0839   -0.82    0.410
## Year2010          -0.0788      0.0879   -0.90    0.371
## Year2011          -0.1598      0.0908   -1.76    0.079 .
## Year2012          -0.1001      0.0906   -1.11    0.270
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.0175, Adjusted R-squared:  -0.0105
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 541 is an outlier with |weight| = 0 ( < 0.00015);
## 55 weights are ~= 1. The remaining 594 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.207  0.861   0.951   0.892   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          1.54e-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 1.226 1           1.107
## Year              1.226 16           1.006

```

Residuals from first author



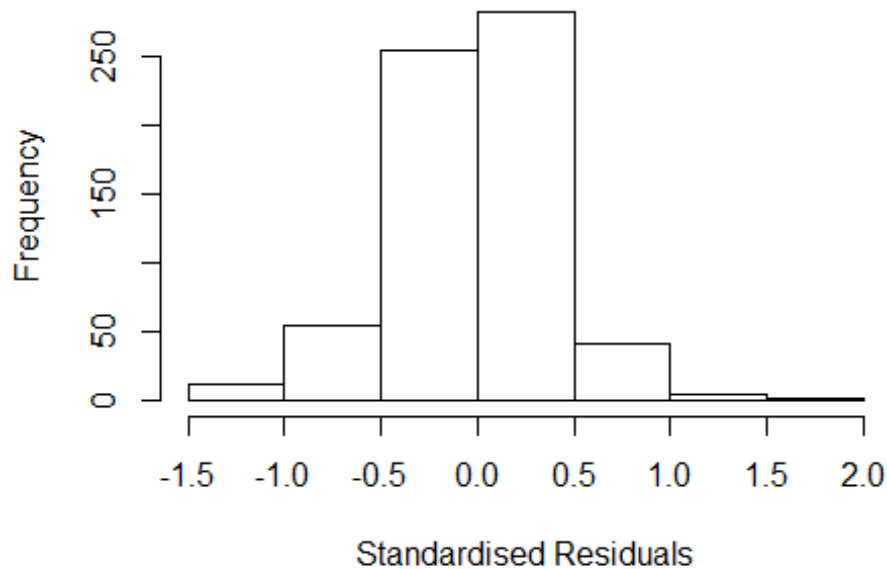
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.21989 -0.22567 0.00877 0.24714 1.93065
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1787 0.0674 17.49 <2e-16 ***
## FirstAuthorFemale1 0.0240 0.0317 0.75 0.451
## Year1997 0.0172 0.0880 0.20 0.845
## Year1998 -0.1708 0.0826 -2.07 0.039 *
## Year1999 -0.1222 0.0937 -1.30 0.193
## Year2000 -0.2071 0.0970 -2.14 0.033 *
## Year2001 -0.1236 0.1488 -0.83 0.407
## Year2002 -0.1243 0.0810 -1.54 0.125
## Year2003 -0.1160 0.0786 -1.48 0.141
## Year2004 -0.0530 0.0961 -0.55 0.582
## Year2005 -0.0514 0.0821 -0.63 0.531
## Year2006 -0.0871 0.0805 -1.08 0.280
```

```

## Year2007          -0.0246      0.0831   -0.30    0.767
## Year2008          -0.0742      0.0812   -0.91    0.361
## Year2009          -0.0704      0.0837   -0.84    0.401
## Year2010          -0.0764      0.0878   -0.87    0.385
## Year2011          -0.1577      0.0906   -1.74    0.082 .
## Year2012          -0.0991      0.0906   -1.09    0.275
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.354
## Multiple R-squared:  0.0173, Adjusted R-squared:  -0.00914
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 541 is an outlier with |weight| = 0 ( < 0.00015);
## 51 weights are ~= 1. The remaining 598 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.197  0.861  0.951  0.892  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      1.54e-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 1.13 1          1.063
## Year              1.13 16          1.004

```

Residuals from last author



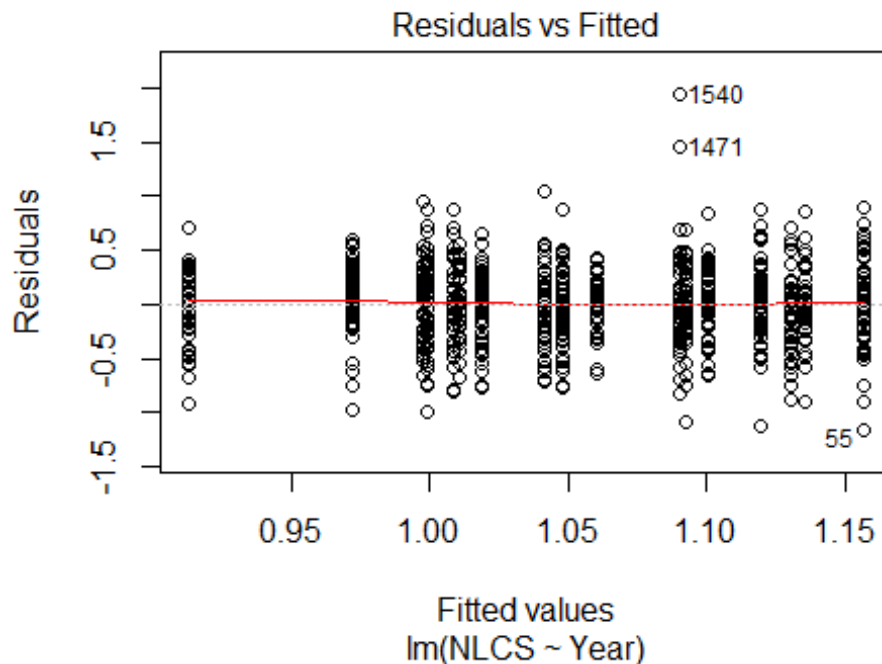
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20213 -0.22981 0.00945 0.24460 1.92171
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1852 0.0658 18.02 <2e-16 ***
## LastAuthorFemale1 0.0169 0.0366 0.46 0.644
## Year1997 0.0147 0.0878 0.17 0.867
## Year1998 -0.1669 0.0844 -1.98 0.049 *
## Year1999 -0.1301 0.0920 -1.41 0.158
## Year2000 -0.1940 0.0945 -2.05 0.041 *
## Year2001 -0.1279 0.1478 -0.87 0.387
## Year2002 -0.1268 0.0806 -1.57 0.116
## Year2003 -0.1184 0.0780 -1.52 0.129
## Year2004 -0.0590 0.0964 -0.61 0.541
## Year2005 -0.0495 0.0827 -0.60 0.550
## Year2006 -0.0872 0.0806 -1.08 0.280
```

```

## Year2007          -0.0260      0.0828   -0.31    0.753
## Year2008          -0.0717      0.0809   -0.89    0.376
## Year2009          -0.0679      0.0840   -0.81    0.419
## Year2010          -0.0739      0.0880   -0.84    0.401
## Year2011          -0.1563      0.0910   -1.72    0.086 .
## Year2012          -0.0969      0.0906   -1.07    0.285
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.0167, Adjusted R-squared:  -0.00977
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 541 is an outlier with |weight| = 0 ( < 0.00015);
## 53 weights are ~= 1. The remaining 596 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.220  0.859  0.950  0.892  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      1.54e-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:  650"
## [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 2008 2009 2010
##   92   80   91   74   93   70   60   52   58   71   76   79   68   92   90
## 2011 2012
##   90   78
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   54   48   45   41   44   37   37   33   47   52   52   54   42   64   71

```

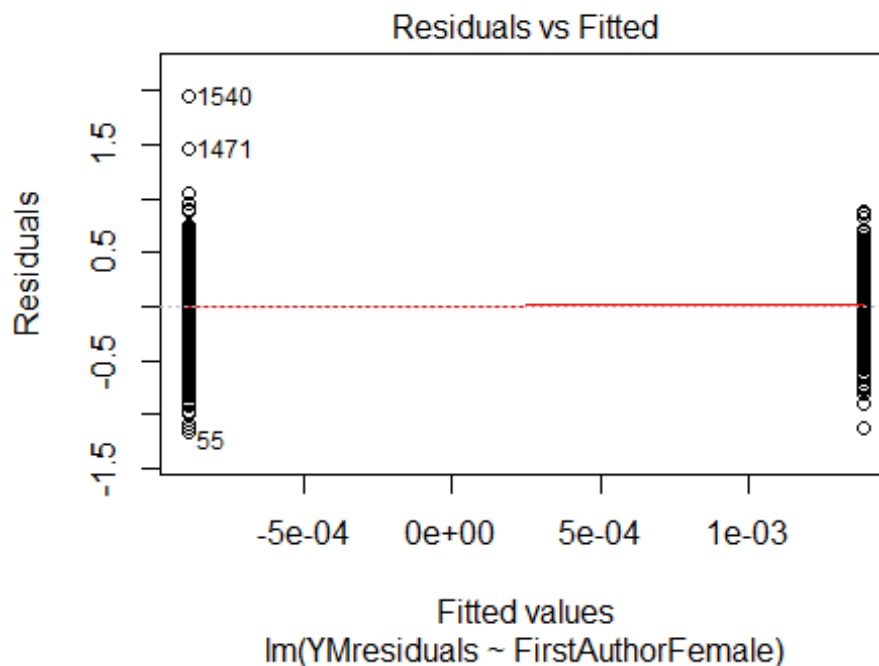
```
## 2011 2012
## 64 59
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 49 43 43 39 40 33 34 30 42 48 42 46 38 58 64
## 2011 2012
## 62 56
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## 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 = 6.6, df = 1, p-value = 0.01
## [1] "Female first author team size 2018 geometric mean: 4.10679628999838"
## [1] "Male first author team size 2018 geometric mean: 3.75424146366177"
## 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 = 230, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.11847494734486"
## [1] "Male last author team size 2018 geometric mean: 3.785096165891"

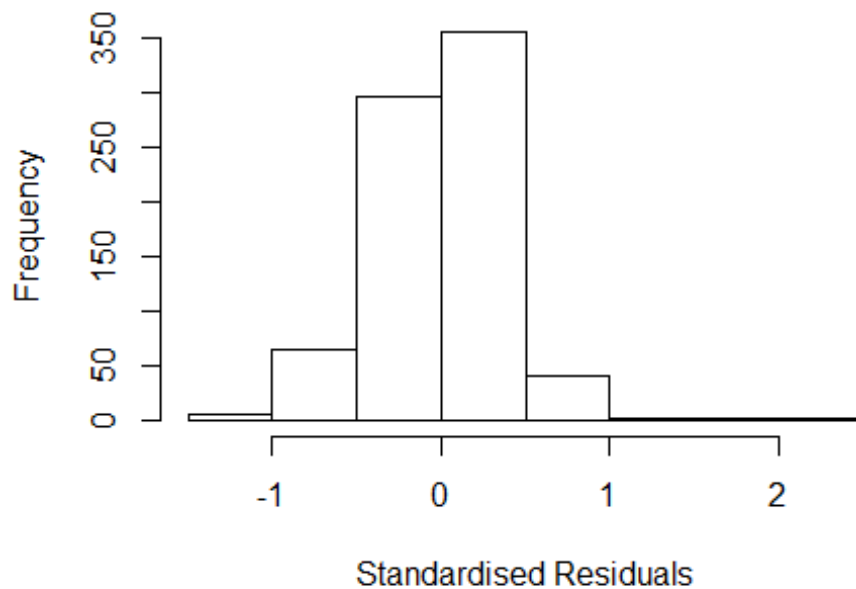
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 220, 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	1.225	1	1.107
LastAuthorFemale	1.093	1	1.045
UniqueAuthors	1.509	4	1.053
Year	1.721	16	1.017

Residuals from first and last author and team size



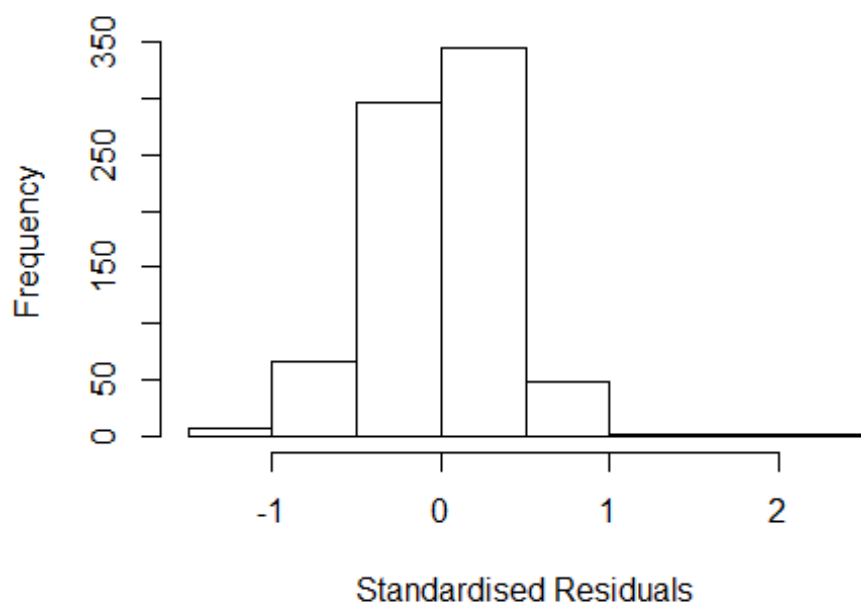
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.229 -0.226 0.017 0.235 2.019
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.054289 0.068728 15.34 < 2e-16 ***
## FirstAuthorFemale1 0.000727 0.027654 0.03 0.97904
## LastAuthorFemale1 -0.046214 0.030710 -1.50 0.13278
## UniqueAuthors2 0.174342 0.057792 3.02 0.00264 **
## UniqueAuthors3 0.178771 0.055904 3.20 0.00144 **
## UniqueAuthors4 0.193458 0.056795 3.41 0.00069 ***
## UniqueAuthors5 0.223282 0.056685 3.94 9e-05 ***
## Year1997 -0.103837 0.098329 -1.06 0.29130
## Year1998 -0.095794 0.089753 -1.07 0.28618
## Year1999 -0.195597 0.080671 -2.42 0.01556 *
```

```

## Year2000      -0.280875    0.083071    -3.38    0.00076 ***
## Year2001      -0.208472    0.098137    -2.12    0.03397 *
## Year2002      -0.233769    0.096509    -2.42    0.01566 *
## Year2003      -0.151176    0.081518    -1.85    0.06406 .
## Year2004      -0.178299    0.093594    -1.91    0.05716 .
## Year2005      -0.139057    0.083404    -1.67    0.09588 .
## Year2006      -0.073160    0.082253    -0.89    0.37405
## Year2007      -0.187722    0.081789    -2.30    0.02200 *
## Year2008      -0.089098    0.085126    -1.05    0.29559
## Year2009      -0.189261    0.078403    -2.41    0.01602 *
## Year2010      -0.214788    0.079341    -2.71    0.00694 **
## Year2011      -0.203088    0.085681    -2.37    0.01803 *
## Year2012      -0.172355    0.080952    -2.13    0.03357 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.0534, Adjusted R-squared:  0.0254
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## observation 637 is an outlier with |weight| = 0 ( < 0.00013);
## 67 weights are ~= 1. The remaining 699 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0019 0.8700 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           1.30e-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 1.162 1 1.078
## LastAuthorFemale 1.086 1 1.042
## Year 1.249 16 1.007

```

Residuals from first and last author

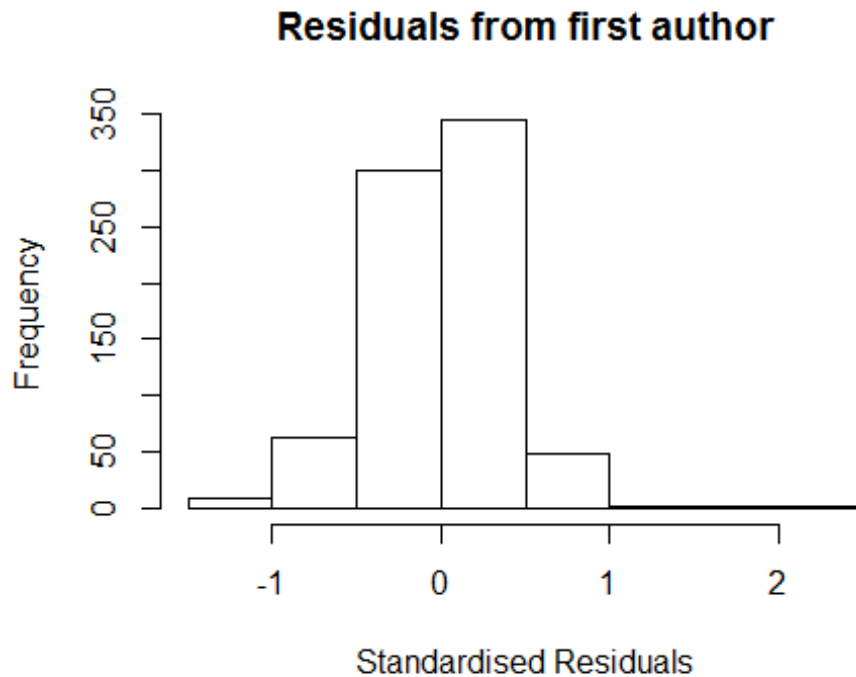


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.162 -0.222 0.012 0.229 2.010
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.16243 0.06676 17.41 <2e-16 ***
## FirstAuthorFemale1 0.02063 0.02717 0.76 0.4479
## LastAuthorFemale1 -0.03714 0.03098 -1.20 0.2310
## Year1997 -0.04260 0.09277 -0.46 0.6462
## Year1998 -0.03897 0.09111 -0.43 0.6690
## Year1999 -0.13340 0.08025 -1.66 0.0969 .
## Year2000 -0.23148 0.08516 -2.72 0.0067 **
## Year2001 -0.15468 0.09594 -1.61 0.1073
## Year2002 -0.17381 0.09771 -1.78 0.0757 .
## Year2003 -0.09507 0.08094 -1.17 0.2405
## Year2004 -0.11799 0.09177 -1.29 0.1989
## Year2005 -0.06535 0.08124 -0.80 0.4214
```

```

## Year2006      -0.00841    0.08100   -0.10    0.9174
## Year2007      -0.11769    0.07832   -1.50    0.1333
## Year2008      -0.02421    0.08175   -0.30    0.7672
## Year2009      -0.12563    0.07652   -1.64    0.1010
## Year2010      -0.13929    0.07638   -1.82    0.0686 .
## Year2011      -0.12667    0.08331   -1.52    0.1288
## Year2012      -0.10319    0.07918   -1.30    0.1929
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.0309, Adjusted R-squared:  0.00758
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## observation 637 is an outlier with |weight| = 0 ( < 0.00013);
## 62 weights are ~= 1. The remaining 704 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0027 0.8710 0.9510 0.8960 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          1.30e-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 1.159 1          1.077
## Year              1.159 16          1.005

```

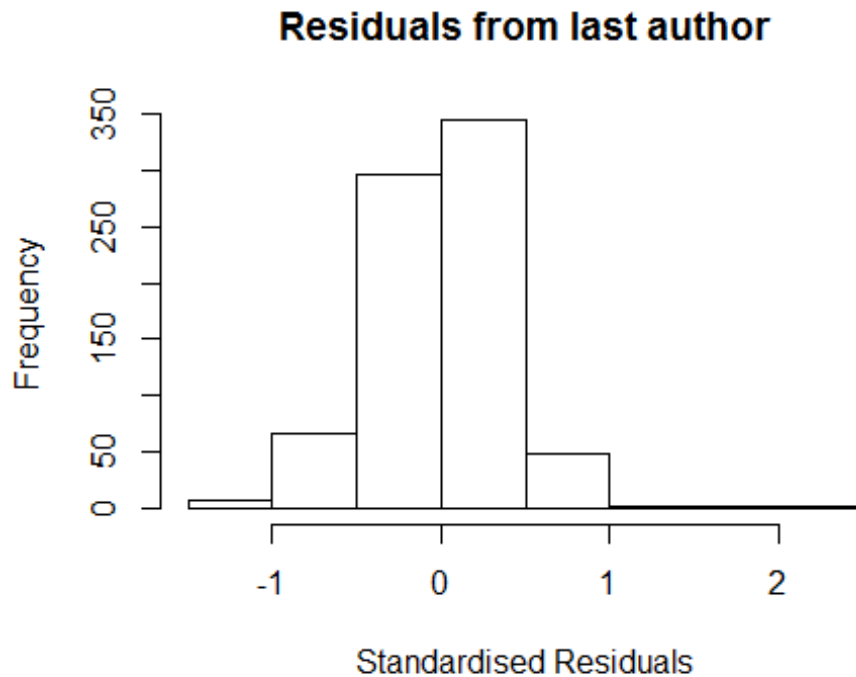


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1568 -0.2220 0.0175 0.2302 2.0179
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1568 0.0662 17.47 <2e-16 ***
## FirstAuthorFemale1 0.0186 0.0272 0.68 0.4936
## Year1997 -0.0406 0.0925 -0.44 0.6605
## Year1998 -0.0376 0.0907 -0.41 0.6790
## Year1999 -0.1294 0.0800 -1.62 0.1060
## Year2000 -0.2333 0.0854 -2.73 0.0065 **
## Year2001 -0.1529 0.0959 -1.59 0.1111
## Year2002 -0.1772 0.0970 -1.83 0.0681 .
## Year2003 -0.0960 0.0803 -1.20 0.2323
## Year2004 -0.1189 0.0907 -1.31 0.1899
## Year2005 -0.0677 0.0812 -0.83 0.4048
## Year2006 -0.0101 0.0811 -0.12 0.9012
```

```

## Year2007          -0.1188      0.0776   -1.53    0.1262
## Year2008          -0.0258      0.0814   -0.32    0.7513
## Year2009          -0.1281      0.0761   -1.68    0.0924 .
## Year2010          -0.1417      0.0756   -1.87    0.0614 .
## Year2011          -0.1354      0.0829   -1.63    0.1028
## Year2012          -0.1085      0.0791   -1.37    0.1708
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.336
## Multiple R-squared:  0.0292, Adjusted R-squared:  0.00715
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## observation 637 is an outlier with |weight| = 0 ( < 0.00013);
## 64 weights are ~= 1. The remaining 702 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0019 0.8690 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      1.30e-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 1.085 1          1.042
## 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1699 -0.2191 0.0168 0.2287 1.9998
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.16986 0.06579 17.78 <2e-16 ***
## LastAuthorFemale1 -0.03569 0.03096 -1.15 0.2494
## Year1997 -0.04626 0.09248 -0.50 0.6170
## Year1998 -0.03952 0.09130 -0.43 0.6653
## Year1999 -0.13530 0.08065 -1.68 0.0938 .
## Year2000 -0.23033 0.08565 -2.69 0.0073 **
## Year2001 -0.15703 0.09544 -1.65 0.1003
## Year2002 -0.17170 0.09818 -1.75 0.0807 .
## Year2003 -0.09339 0.08148 -1.15 0.2521
## Year2004 -0.12178 0.09193 -1.32 0.1857
## Year2005 -0.06571 0.08158 -0.81 0.4208
## Year2006 -0.00839 0.08169 -0.10 0.9183
```

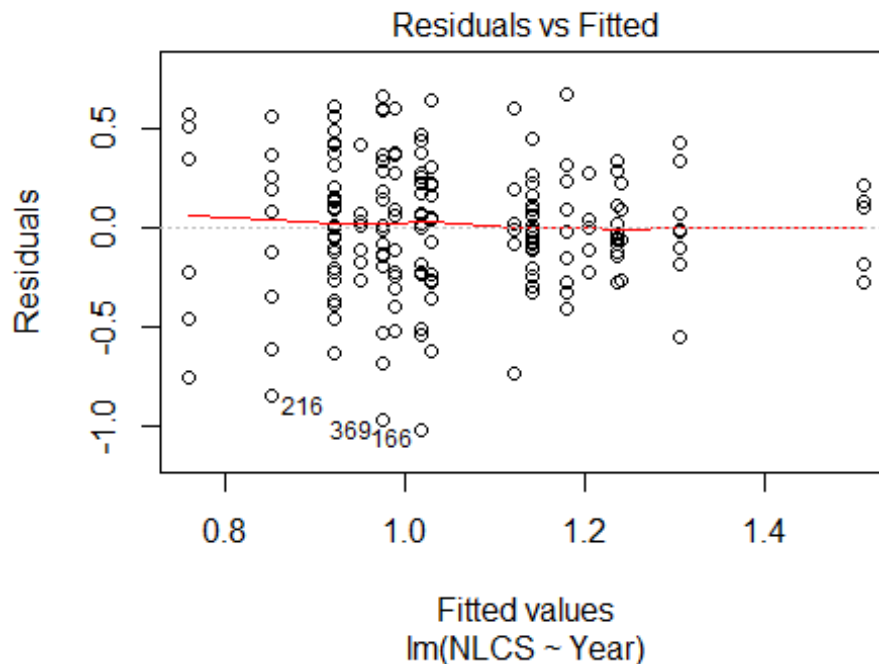
```

## Year2007          -0.11909      0.07879    -1.51    0.1311
## Year2008          -0.02025      0.08263    -0.25    0.8065
## Year2009          -0.12442      0.07715    -1.61    0.1073
## Year2010          -0.13669      0.07725    -1.77    0.0772 .
## Year2011          -0.12433      0.08406    -1.48    0.1395
## Year2012          -0.09924      0.07974    -1.24    0.2137
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.0299, Adjusted R-squared:  0.00784
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## observation 637 is an outlier with |weight| = 0 ( < 0.00013);
## 60 weights are ~= 1. The remaining 706 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0058 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          1.30e-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: 767"
## [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
##    10    6    6   18   12   17    9   12   16   21   16   16   15   18   14
## 2011 2012
##    29   40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    4   12   10    5    6    8   12   19   11   12   12   14    7

```



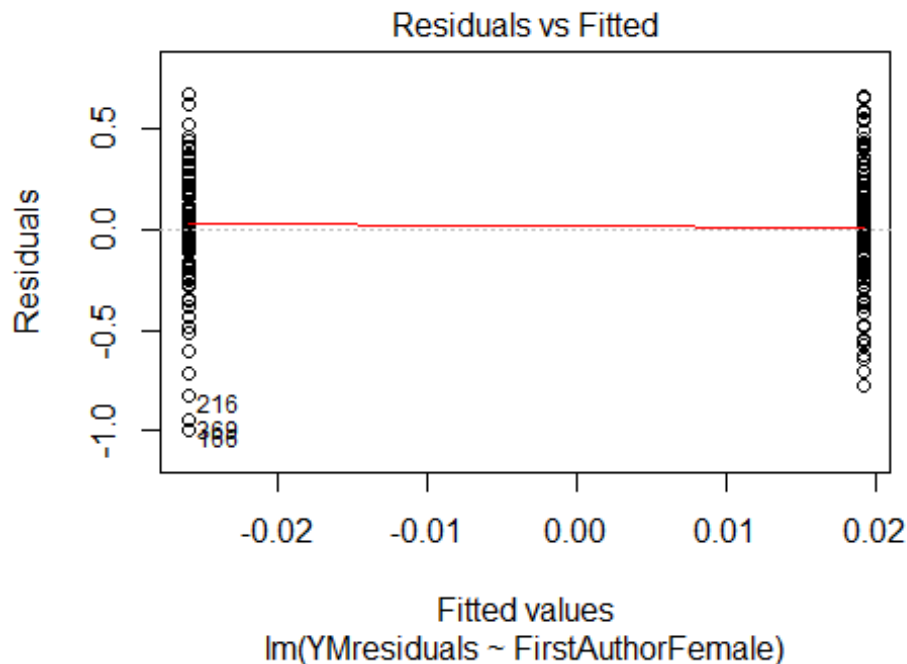
```
## 2011 2012
## 21 36
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 4 4 12 10 4 6 8 11 19 10 11 11 14 7
## 2011 2012
## 19 32
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## 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 = 0.003, df = 1, p-value = 1
## [1] "Female first author team size 2018 geometric mean: 5.10791064855846"
## [1] "Male first author team size 2018 geometric mean: 4.277128053329"
## 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 = 36, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.45963911711622"
## [1] "Male last author team size 2018 geometric mean: 4.74288121955862"

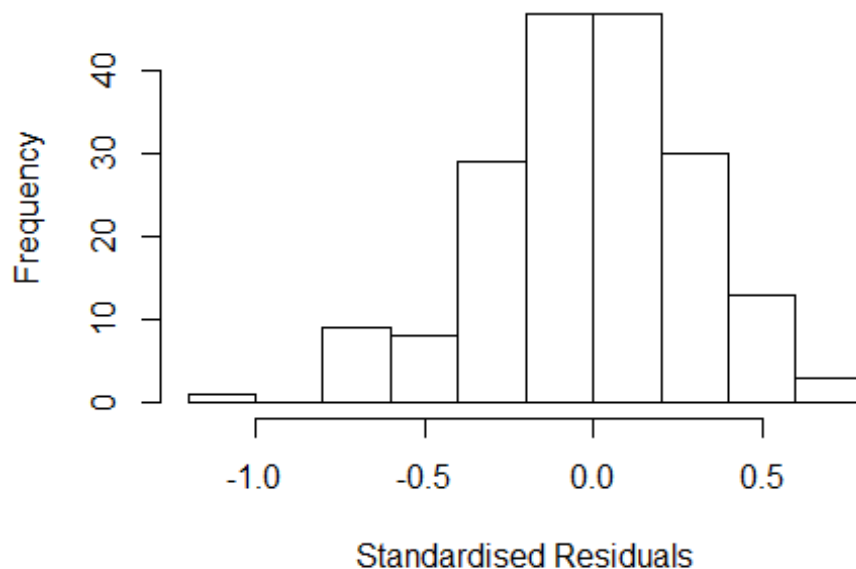
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 30, 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	2.044	1	1.430
LastAuthorFemale	1.895	1	1.376
UniqueAuthors	5.443	4	1.236
Year	15.217	16	1.089

Residuals from first and last author and team size



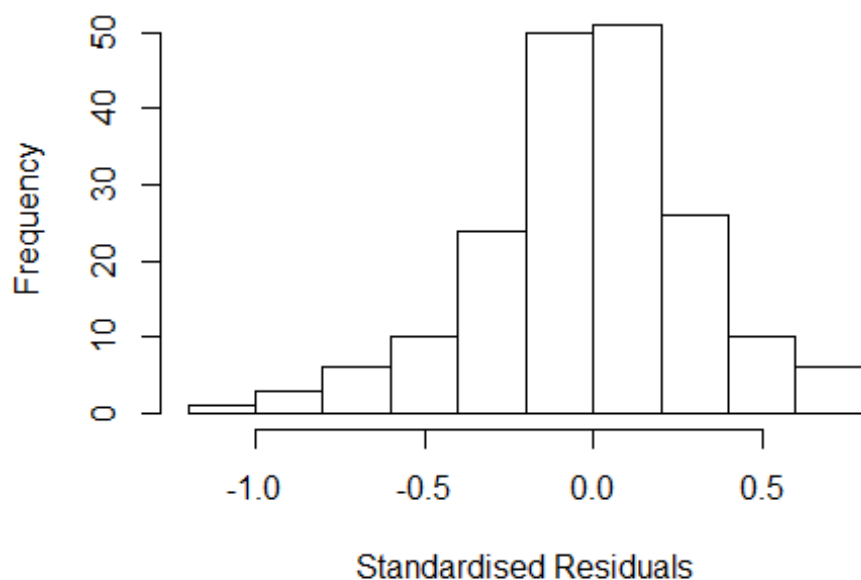
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.07448 -0.19906 -0.00224 0.19391 0.61156
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.74683 0.27128 2.75 0.0066 **
## FirstAuthorFemale1 -0.02389 0.05358 -0.45 0.6563
## LastAuthorFemale1 -0.06859 0.06909 -0.99 0.3223
## UniqueAuthors2 0.16832 0.14123 1.19 0.2350
## UniqueAuthors3 0.19859 0.13679 1.45 0.1485
## UniqueAuthors4 0.29536 0.14320 2.06 0.0407 *
## UniqueAuthors5 0.32765 0.12647 2.59 0.0104 *
## Year1997 0.53820 0.26473 2.03 0.0437 *
## Year1998 0.30252 0.25427 1.19 0.2359
## Year1999 0.16238 0.24278 0.67 0.5045
```

```

## Year2000      0.20125    0.24637    0.82    0.4152
## Year2001      0.23919    0.22779    1.05    0.2952
## Year2002      0.17104    0.26202    0.65    0.5148
## Year2003      0.27572    0.26540    1.04    0.3004
## Year2004      0.00765    0.27117    0.03    0.9775
## Year2005      0.08740    0.23655    0.37    0.7123
## Year2006     -0.06075    0.24763   -0.25    0.8065
## Year2007      0.15631    0.23491    0.67    0.5067
## Year2008      0.14535    0.23896    0.61    0.5439
## Year2009      0.03999    0.25157    0.16    0.8739
## Year2010     -0.01275    0.24566   -0.05    0.9587
## Year2011      0.00972    0.25866    0.04    0.9701
## Year2012      0.00166    0.24248    0.01    0.9946
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.303
## Multiple R-squared:  0.226, Adjusted R-squared:  0.122
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 170 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.184  0.883  0.953  0.904  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      5.35e-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 1.754 1      1.325
## LastAuthorFemale  1.519 1      1.233
## Year              2.661 16      1.031

```

Residuals from first and last author



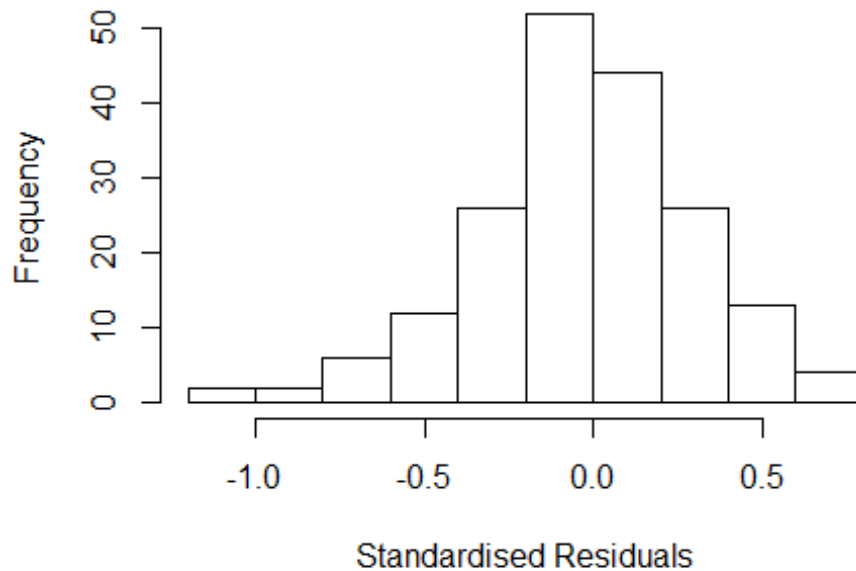
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.01243 -0.18548 -0.00888 0.18001 0.73817
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.012430 0.290636 3.48 0.00063 ***
## FirstAuthorFemale1 -0.010082 0.052980 -0.19 0.84930
## LastAuthorFemale1 -0.121038 0.068274 -1.77 0.07807 .
## Year1997 0.479448 0.304417 1.57 0.11714
## Year1998 0.257098 0.300735 0.85 0.39382
## Year1999 0.141198 0.297925 0.47 0.63616
## Year2000 0.157049 0.314044 0.50 0.61767
## Year2001 0.244553 0.300316 0.81 0.41661
## Year2002 0.146576 0.330754 0.44 0.65822
## Year2003 0.321345 0.312524 1.03 0.30532
## Year2004 -0.036558 0.322829 -0.11 0.90997
## Year2005 0.096126 0.296963 0.32 0.74657
```

```

## Year2006      -0.029724    0.319696   -0.09   0.92603
## Year2007      0.217595    0.295822    0.74   0.46302
## Year2008      0.204331    0.300553    0.68   0.49753
## Year2009      0.032188    0.303143    0.11   0.91557
## Year2010     -0.001265    0.310036    0.00   0.99675
## Year2011      0.018568    0.311858    0.06   0.95259
## Year2012     -0.000984    0.302345    0.00   0.99741
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.305
## Multiple R-squared:  0.159, Adjusted R-squared:  0.0693
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 174 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.248  0.880   0.962   0.900   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      5.35e-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 1.653 1      1.286
## Year              1.653 16      1.016

```

Residuals from first author



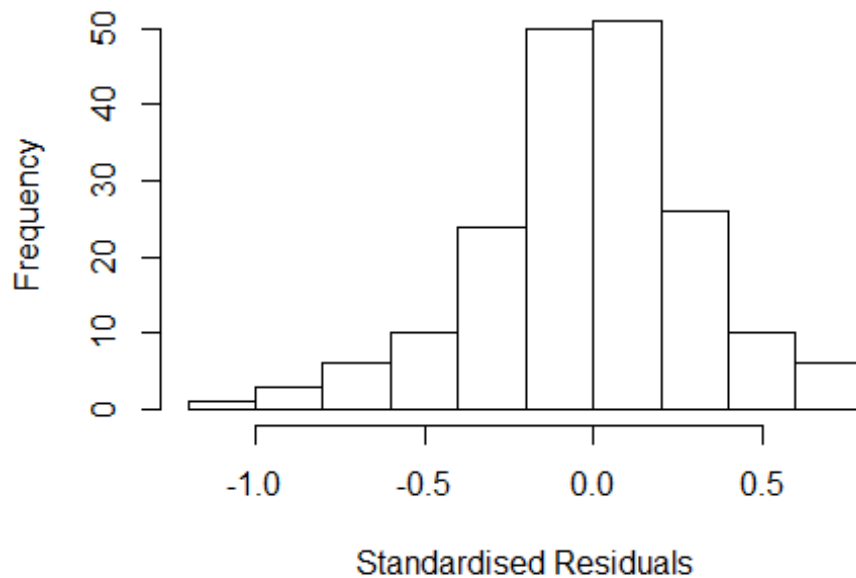
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06710 -0.21260 -0.00967 0.17294 0.69692
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0168 0.2904 3.50 0.00059 ***
## FirstAuthorFemale1 -0.0209 0.0539 -0.39 0.69886
## Year1997 0.4481 0.3067 1.46 0.14579
## Year1998 0.2247 0.3050 0.74 0.46228
## Year1999 0.1258 0.2970 0.42 0.67231
## Year2000 0.1413 0.3161 0.45 0.65552
## Year2001 0.2401 0.3002 0.80 0.42485
## Year2002 0.1558 0.3313 0.47 0.63875
## Year2003 0.3027 0.3109 0.97 0.33165
## Year2004 -0.0627 0.3237 -0.19 0.84675
## Year2005 0.0712 0.2966 0.24 0.81061
## Year2006 -0.0522 0.3186 -0.16 0.87001
```

```

## Year2007          0.1915      0.2950      0.65  0.51703
## Year2008          0.1600      0.2968      0.54  0.59047
## Year2009          0.0244      0.3020      0.08  0.93578
## Year2010         -0.0675      0.3029     -0.22  0.82390
## Year2011         -0.0117      0.3109     -0.04  0.97000
## Year2012         -0.0471      0.3009     -0.16  0.87569
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.303
## Multiple R-squared:  0.136, Adjusted R-squared:  0.0495
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 164 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.188  0.853   0.948   0.889   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      5.35e-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 1.516 1          1.231
## Year            1.516 16          1.013

```


Residuals from last author



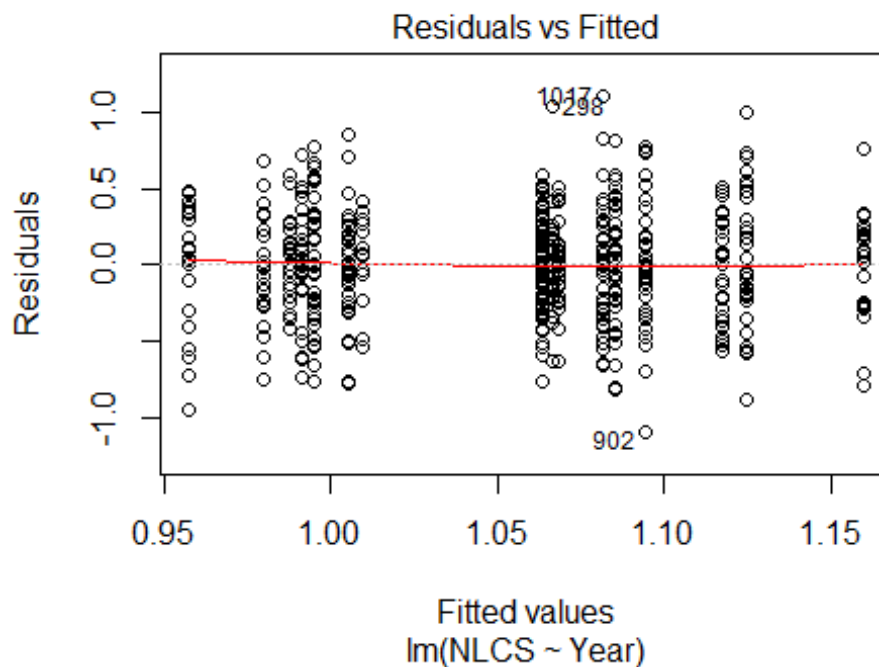
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.01361 -0.19065 -0.00731  0.17588  0.74511
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.01361    0.29008   3.49  0.00061 ***
## LastAuthorFemale1 -0.12216    0.06890  -1.77  0.07803 .
## Year1997         0.47382    0.30285   1.56  0.11957
## Year1998         0.25619    0.30026   0.85  0.39475
## Year1999         0.13818    0.29716   0.47  0.64252
## Year2000         0.15265    0.31248   0.49  0.62582
## Year2001         0.24336    0.29977   0.81  0.41803
## Year2002         0.13808    0.32701   0.42  0.67338
## Year2003         0.31645    0.31104   1.02  0.31042
## Year2004        -0.04356    0.31795  -0.14  0.89118
## Year2005         0.08970    0.29534   0.30  0.76172
## Year2006        -0.03281    0.31907  -0.10  0.91823
```

```

## Year2007      0.21503      0.29506      0.73  0.46715
## Year2008      0.19877      0.29887      0.67  0.50691
## Year2009      0.02598      0.30210      0.09  0.93156
## Year2010     -0.00499      0.30861     -0.02  0.98711
## Year2011      0.01404      0.31014      0.05  0.96394
## Year2012     -0.00811      0.29828     -0.03  0.97834
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.304
## Multiple R-squared:  0.159, Adjusted R-squared:  0.0749
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 174 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.245  0.879  0.962  0.899  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      5.35e-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: 187"
## [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
##   53   61   51   51   46   50   64   38   42   46   49   60   41   54   53
## 2011 2012
##   68   64
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   30   26   24   22   22   15   41   25   26   29   39   39   24   36   37
## 2011 2012

```

```
## 45 42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 28 22 23 21 16 12 31 21 24 22 32 32 21 33 33
## 2011 2012
## 41 38
## [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
```



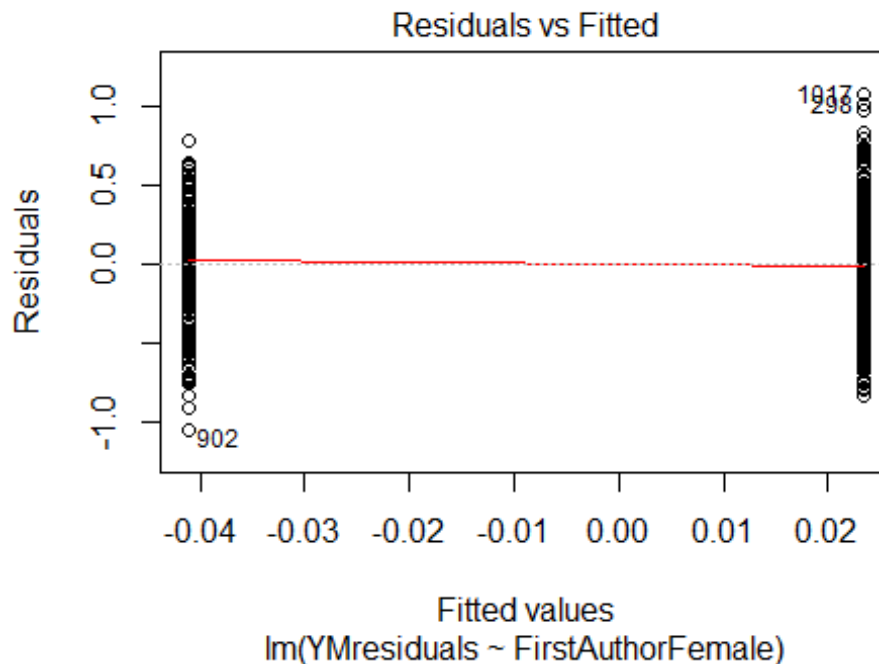
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0088, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 3.92304536876973"
## [1] "Male first author team size 2018 geometric mean: 3.62688337482331"

## 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: 2.85009967337652"
## [1] "Male last author team size 2018 geometric mean: 4.31312072606741"

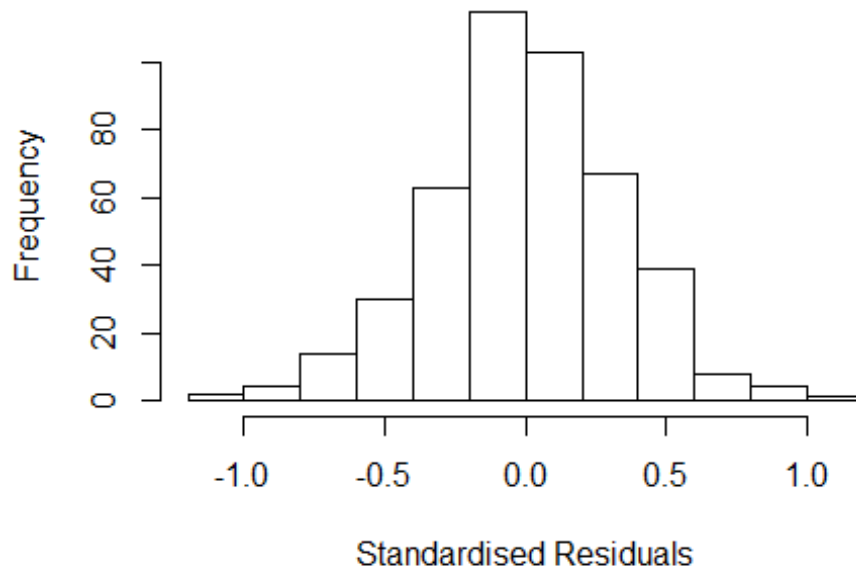
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 72, 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	1.355	1	1.164
LastAuthorFemale	1.235	1	1.111
UniqueAuthors	2.012	4	1.091
Year	2.556	16	1.030

Residuals from first and last author and team size



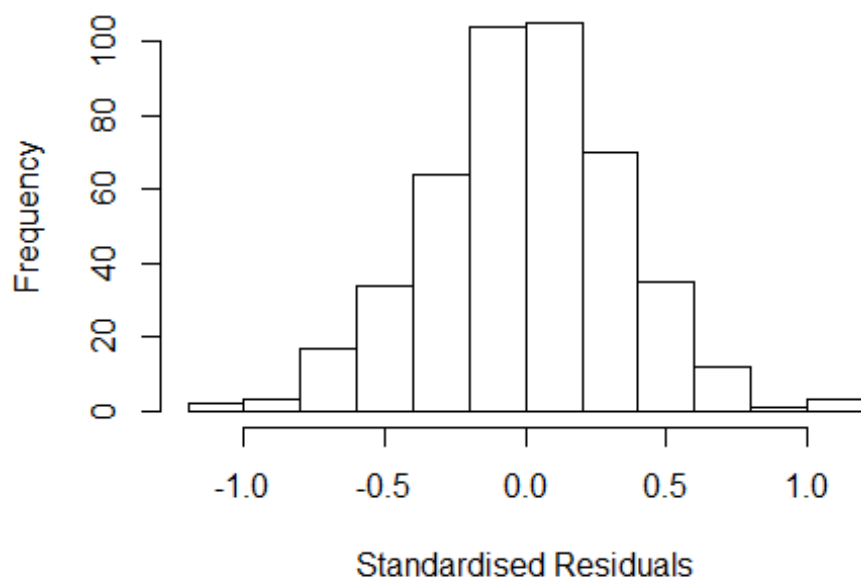
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.16003 -0.20088 -0.00702 0.22079 1.01585
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9545 0.1082 8.82 <2e-16 ***
## FirstAuthorFemale1 -0.0482 0.0367 -1.31 0.1894
## LastAuthorFemale1 -0.0744 0.0390 -1.91 0.0573 .
## UniqueAuthors2 0.0917 0.0911 1.01 0.3146
## UniqueAuthors3 0.1270 0.0910 1.40 0.1636
## UniqueAuthors4 0.2634 0.0892 2.95 0.0033 **
## UniqueAuthors5 0.3098 0.0896 3.46 0.0006 ***
## Year1997 -0.1454 0.1088 -1.34 0.1824
## Year1998 -0.1048 0.0923 -1.14 0.2568
## Year1999 -0.0217 0.1122 -0.19 0.8468
```

```

## Year2000          -0.0417      0.1033   -0.40   0.6871
## Year2001          -0.0869      0.1268   -0.68   0.4939
## Year2002          -0.0938      0.1039   -0.90   0.3675
## Year2003           0.0137      0.1112    0.12   0.9021
## Year2004          -0.0262      0.0966   -0.27   0.7861
## Year2005          -0.0936      0.0903   -1.04   0.3002
## Year2006          -0.0524      0.1037   -0.51   0.6136
## Year2007          -0.1347      0.1019   -1.32   0.1867
## Year2008          -0.1350      0.1107   -1.22   0.2234
## Year2009          -0.0561      0.1023   -0.55   0.5838
## Year2010          -0.0467      0.1054   -0.44   0.6578
## Year2011          -0.0506      0.0963   -0.53   0.5997
## Year2012           0.0182      0.0947    0.19   0.8478
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.319
## Multiple R-squared:  0.111, Adjusted R-squared:  0.0647
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 414 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.159  0.864  0.952  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      2.22e-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 1.389 1      1.178
## LastAuthorFemale  1.161 1      1.077
## Year              1.564 16      1.014

```

Residuals from first and last author



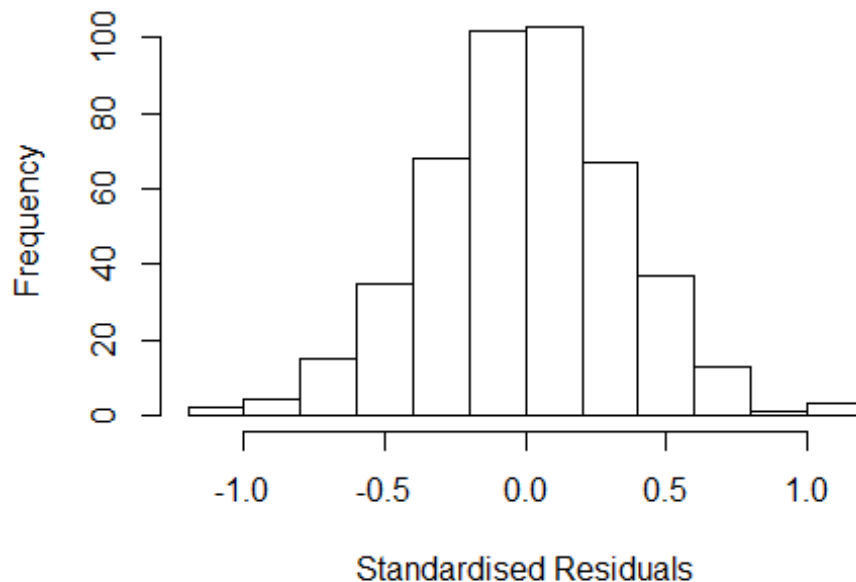
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.09179 -0.22336 0.00597 0.22822 1.07008
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.11360 0.07586 14.68 <2e-16 ***
## FirstAuthorFemale1 -0.04910 0.03859 -1.27 0.20
## LastAuthorFemale1 -0.04173 0.04022 -1.04 0.30
## Year1997 -0.15190 0.11028 -1.38 0.17
## Year1998 -0.08224 0.09537 -0.86 0.39
## Year1999 -0.03667 0.11404 -0.32 0.75
## Year2000 -0.02356 0.09784 -0.24 0.81
## Year2001 -0.08849 0.12809 -0.69 0.49
## Year2002 -0.06464 0.10437 -0.62 0.54
## Year2003 0.06178 0.10173 0.61 0.54
## Year2004 -0.01303 0.09716 -0.13 0.89
## Year2005 -0.04682 0.08892 -0.53 0.60
```

```

## Year2006      -0.02607    0.10221   -0.26    0.80
## Year2007      -0.06851    0.10071   -0.68    0.50
## Year2008      -0.08562    0.11398   -0.75    0.45
## Year2009       0.02729    0.10467    0.26    0.79
## Year2010       0.00333    0.10812    0.03    0.98
## Year2011      -0.00902    0.09421   -0.10    0.92
## Year2012       0.03469    0.09774    0.35    0.72
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.333
## Multiple R-squared:  0.0286, Adjusted R-squared:  -0.0119
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 42 weights are ~= 1. The remaining 408 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.259  0.863   0.948   0.899   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      2.22e-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 1.361 1      1.166
## Year              1.361 16      1.010

```


Residuals from first author



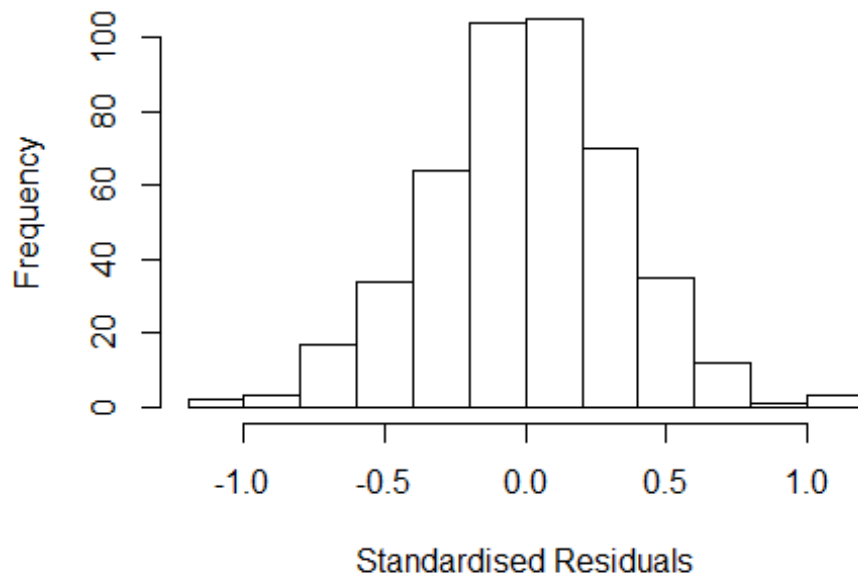
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.08229 -0.22760 -0.00196 0.22009 1.08241
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10963 0.07575 14.65 <2e-16 ***
## FirstAuthorFemale1 -0.05238 0.03824 -1.37 0.17
## Year1997 -0.14984 0.11045 -1.36 0.18
## Year1998 -0.08686 0.09531 -0.91 0.36
## Year1999 -0.04444 0.11363 -0.39 0.70
## Year2000 -0.02380 0.09819 -0.24 0.81
## Year2001 -0.08269 0.12782 -0.65 0.52
## Year2002 -0.06971 0.10369 -0.67 0.50
## Year2003 0.06384 0.10137 0.63 0.53
## Year2004 -0.02189 0.09628 -0.23 0.82
## Year2005 -0.04746 0.08931 -0.53 0.60
## Year2006 -0.03197 0.10179 -0.31 0.75
```

```

## Year2007          -0.06882    0.10134   -0.68    0.50
## Year2008          -0.08698    0.11295   -0.77    0.44
## Year2009           0.02504    0.10446    0.24    0.81
## Year2010          -0.00503    0.10719   -0.05    0.96
## Year2011          -0.01074    0.09414   -0.11    0.91
## Year2012           0.02705    0.09844    0.27    0.78
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.332
## Multiple R-squared:  0.0266, Adjusted R-squared:  -0.0117
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 43 weights are ~= 1. The remaining 407 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.266  0.864  0.950  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      2.22e-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 1.148 1      1.072
## Year              1.148 16      1.004

```

Residuals from last author



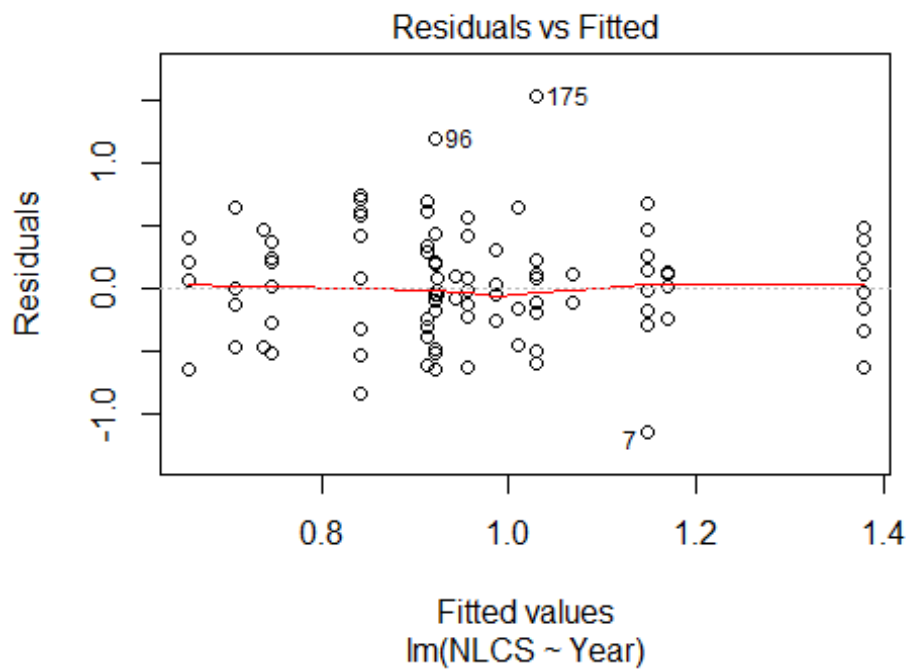
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.12319 -0.21781 -0.00443 0.22376 1.08899
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10619 0.07603 14.55 <2e-16 ***
## LastAuthorFemale1 -0.04726 0.03981 -1.19 0.24
## Year1997 -0.15442 0.11102 -1.39 0.16
## Year1998 -0.09326 0.09488 -0.98 0.33
## Year1999 -0.04424 0.11304 -0.39 0.70
## Year2000 -0.03301 0.09607 -0.34 0.73
## Year2001 -0.10778 0.12990 -0.83 0.41
## Year2002 -0.07344 0.10355 -0.71 0.48
## Year2003 0.06538 0.10071 0.65 0.52
## Year2004 -0.01189 0.09676 -0.12 0.90
## Year2005 -0.06135 0.08819 -0.70 0.49
## Year2006 -0.03010 0.10205 -0.29 0.77
```

```

## Year2007          -0.08381      0.10013    -0.84      0.40
## Year2008          -0.10544      0.11250    -0.94      0.35
## Year2009           0.01700      0.10330     0.16      0.87
## Year2010          -0.00817      0.10659    -0.08      0.94
## Year2011          -0.03211      0.09058    -0.35      0.72
## Year2012           0.02573      0.09703     0.27      0.79
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.332
## Multiple R-squared:  0.0246, Adjusted R-squared:  -0.0138
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 40 weights are ~= 1. The remaining 410 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.230  0.869  0.950  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      2.22e-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: 450"
## [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
##   11   8   5   7   8   6   6   6   4  13  15   9  18   8   9
## 2011 2012
##   10  10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   9   8   3   3   6   2   4   2   4  10  11   4   9   5   9
## 2011 2012

```

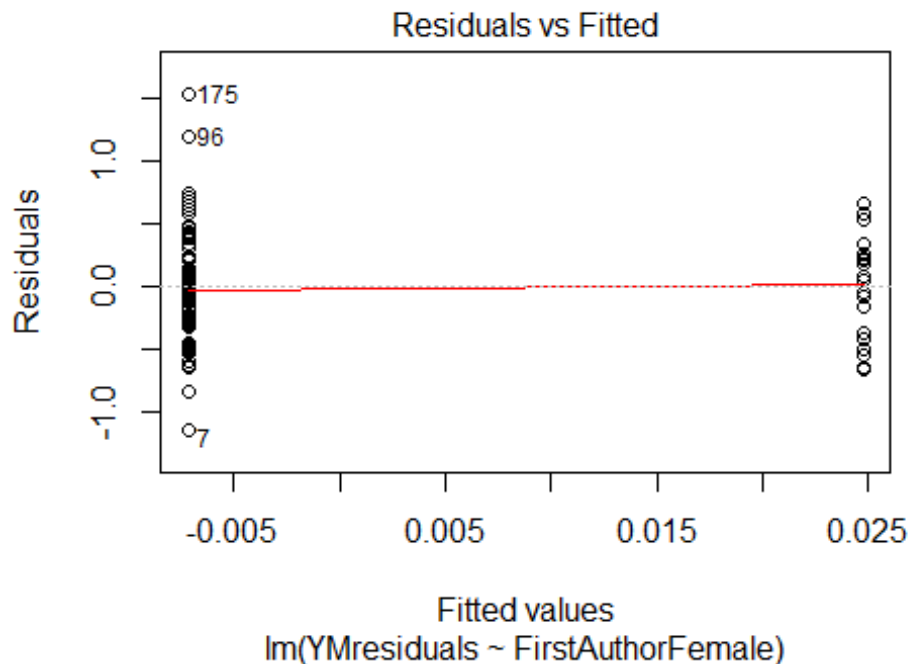
```
##      2      7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      8      8      3      3      6      2      4      2      3      8     11      3      8      4      7
## 2011 2012
##      2      6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.88, df = 1, p-value = 0.3
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 2.71441761659491"
## 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 = 2, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male last author team size 2018 geometric mean: 2.46621207433047"

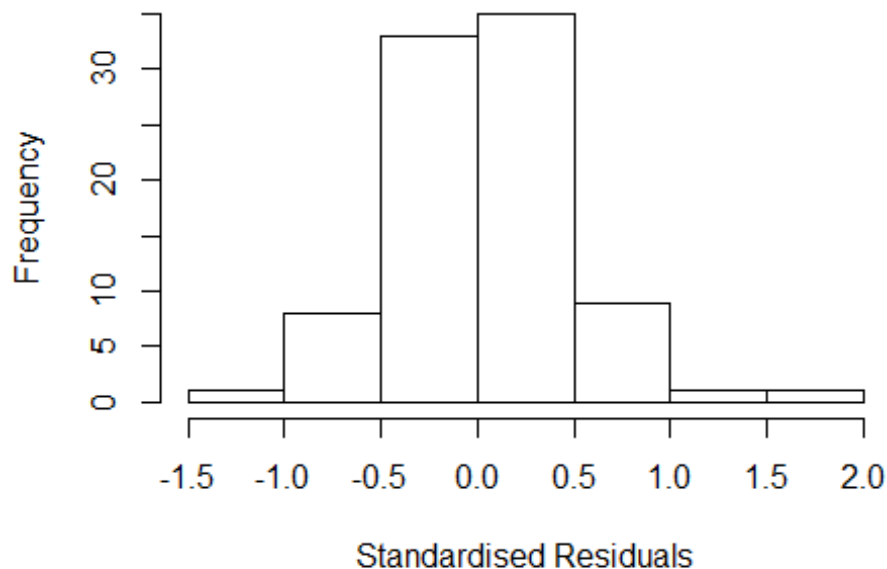
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3.5, 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	4.713	1	2.171
LastAuthorFemale	5.255	1	2.292
UniqueAuthors	78.974	4	1.727
Year	470.779	16	1.212

Residuals from first and last author and team size



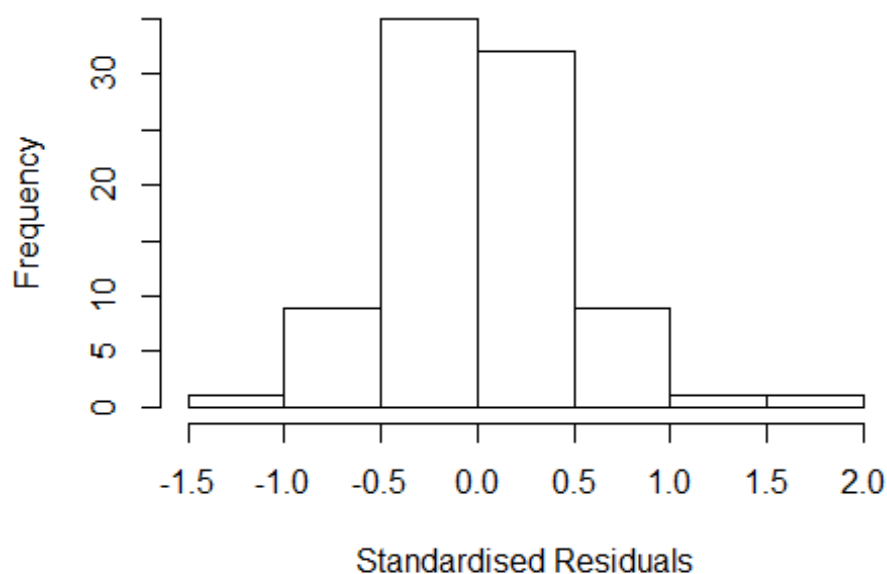
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.04067 -0.23233 0.00258 0.27109 1.77686
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2971 0.1946 6.67 6.7e-09 ***
## FirstAuthorFemale1 0.1853 0.1190 1.56 0.124
## LastAuthorFemale1 -0.2699 0.1976 -1.37 0.177
## UniqueAuthors2 -0.3431 0.2745 -1.25 0.216
## UniqueAuthors3 -0.2318 0.1830 -1.27 0.210
## UniqueAuthors4 -0.2364 0.1821 -1.30 0.199
## UniqueAuthors5 -0.1026 0.2218 -0.46 0.645
## Year1997 0.2890 0.2001 1.44 0.153
## Year1998 -0.1873 0.3974 -0.47 0.639
## Year1999 -0.2010 0.1532 -1.31 0.194
```

```

## Year2000          -0.2828      0.2638   -1.07    0.288
## Year2001          -0.0591      0.1740   -0.34    0.735
## Year2002          -0.1740      0.2143   -0.81    0.420
## Year2003          -0.1190      0.1742   -0.68    0.497
## Year2004          -0.2888      0.3422   -0.84    0.402
## Year2005          -0.1383      0.2307   -0.60    0.551
## Year2006          -0.1538      0.2658   -0.58    0.565
## Year2007          -0.3763      0.4381   -0.86    0.393
## Year2008          -0.3879      0.2155   -1.80    0.076 .
## Year2009           0.2071      0.1713    1.21    0.231
## Year2010          -0.2892      0.2238   -1.29    0.201
## Year2011          -0.4851      0.5076   -0.96    0.343
## Year2012          -0.0849      0.2167   -0.39    0.697
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.259, Adjusted R-squared:  0.00871
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0733 0.8820 0.9540 0.9060 0.9900 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          1.14e-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 2.994 1          1.730
## LastAuthorFemale 1.713 1          1.309
## Year              5.044 16          1.052

```


Residuals from first and last author



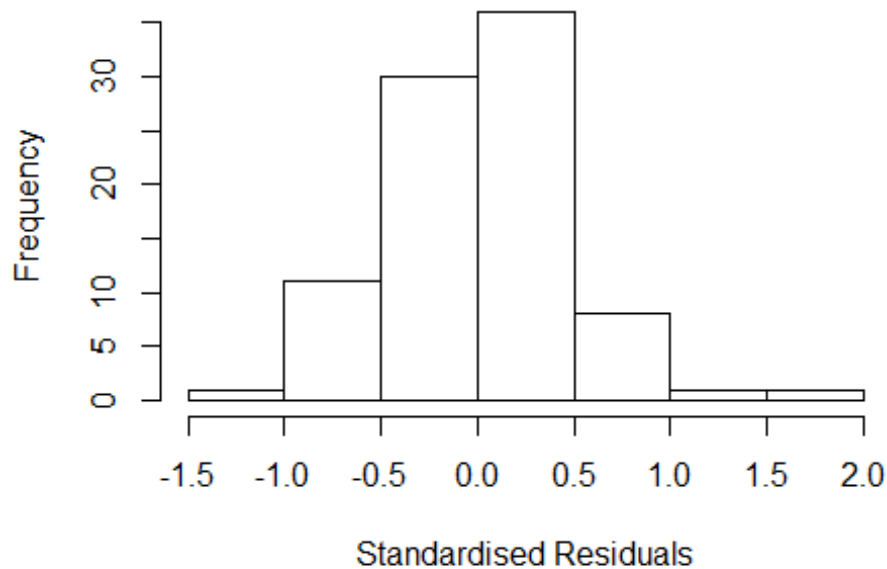
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.14421 -0.24492 -0.00567 0.28940 1.72555
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1442 0.1789 6.40 1.6e-08 ***
## FirstAuthorFemale1 0.1534 0.1197 1.28 0.204
## LastAuthorFemale1 -0.2898 0.1538 -1.88 0.064 .
## Year1997 0.2606 0.2109 1.24 0.221
## Year1998 -0.1657 0.3679 -0.45 0.654
## Year1999 -0.2723 0.1724 -1.58 0.119
## Year2000 -0.3912 0.2116 -1.85 0.069 .
## Year2001 -0.0757 0.1965 -0.39 0.701
## Year2002 -0.1590 0.2074 -0.77 0.446
## Year2003 -0.2002 0.1908 -1.05 0.298
## Year2004 -0.4025 0.3618 -1.11 0.270
## Year2005 -0.1751 0.2230 -0.79 0.435
```

```

## Year2006          -0.1941      0.2795   -0.69    0.490
## Year2007          -0.4498      0.4450   -1.01    0.316
## Year2008          -0.3940      0.2351   -1.68    0.098 .
## Year2009           0.1617      0.1867    0.87    0.390
## Year2010          -0.3168      0.2145   -1.48    0.144
## Year2011          -0.4834      0.5739   -0.84    0.402
## Year2012          -0.1219      0.2269   -0.54    0.593
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.439
## Multiple R-squared:  0.23,   Adjusted R-squared:  0.0286
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0872 0.8700 0.9580 0.9040 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      1.14e-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 4.938 1      2.222
## Year              4.938 16      1.051

```

Residuals from first author



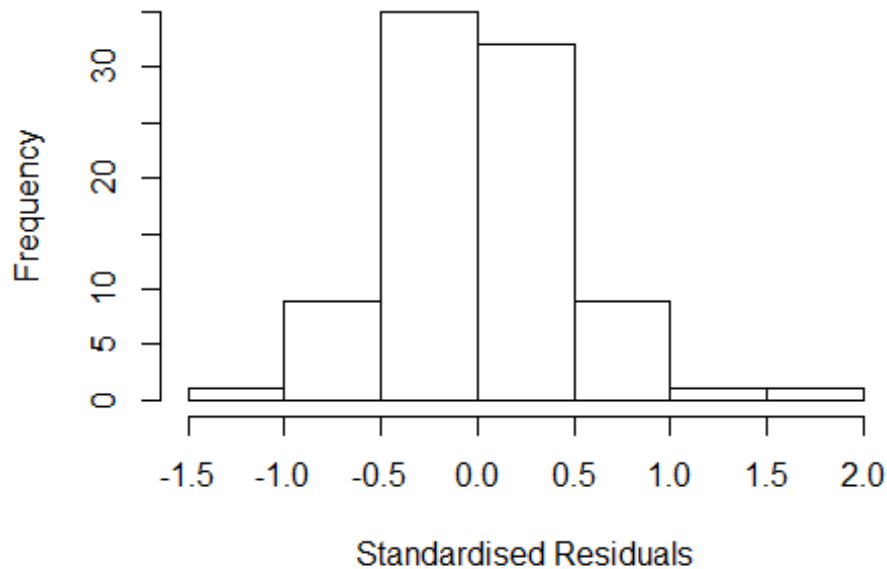
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11158 -0.25969 0.00829 0.27402 1.72102
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1116 0.1682 6.61 6.4e-09 ***
## FirstAuthorFemale1 0.1182 0.1279 0.92 0.358
## Year1997 0.2621 0.2146 1.22 0.226
## Year1998 -0.1338 0.3625 -0.37 0.713
## Year1999 -0.2280 0.1550 -1.47 0.146
## Year2000 -0.3979 0.1927 -2.06 0.043 *
## Year2001 -0.0431 0.1868 -0.23 0.818
## Year2002 -0.1264 0.1982 -0.64 0.526
## Year2003 -0.1676 0.1808 -0.93 0.357
## Year2004 -0.3702 0.3568 -1.04 0.303
## Year2005 -0.2209 0.2473 -0.89 0.375
## Year2006 -0.2603 0.3058 -0.85 0.397
```

```

## Year2007          -0.5021      0.3385    -1.48     0.142
## Year2008          -0.3470      0.2232    -1.55     0.125
## Year2009           0.1222      0.1703     0.72     0.476
## Year2010          -0.2796      0.2024    -1.38     0.172
## Year2011          -0.4332      0.5417    -0.80     0.427
## Year2012          -0.1553      0.2266    -0.69     0.495
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.185, Adjusted R-squared:  -0.0134
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.082  0.870   0.957   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      1.14e-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 1.64 1          1.280
## Year              1.64 16          1.016

```

Residuals from last author



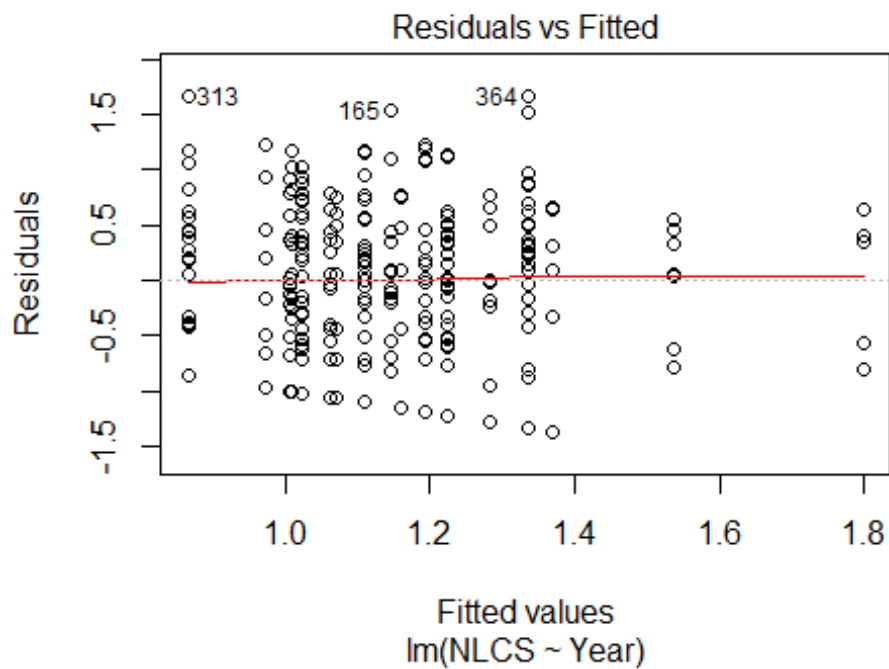
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1870 -0.2676 -0.0125 0.2864 1.6919
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.187 0.175 6.80 2.9e-09 ***
## LastAuthorFemale1 -0.264 0.158 -1.66 0.10
## Year1997 0.235 0.209 1.12 0.26
## Year1998 -0.208 0.366 -0.57 0.57
## Year1999 -0.264 0.178 -1.48 0.14
## Year2000 -0.383 0.223 -1.72 0.09 .
## Year2001 -0.118 0.192 -0.62 0.54
## Year2002 -0.202 0.204 -0.99 0.33
## Year2003 -0.243 0.187 -1.30 0.20
## Year2004 -0.445 0.360 -1.24 0.22
## Year2005 -0.144 0.225 -0.64 0.52
## Year2006 -0.238 0.282 -0.85 0.40
```

```

## Year2007          -0.500          0.429        -1.16          0.25
## Year2008          -0.372          0.247        -1.51          0.14
## Year2009           0.113          0.183          0.62          0.54
## Year2010          -0.326          0.226        -1.44          0.15
## Year2011          -0.449          0.461        -0.97          0.33
## Year2012          -0.101          0.227        -0.44          0.66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.212, Adjusted R-squared:  0.0209
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.113  0.895  0.958  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          1.14e-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: 88"
## [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
##   14   12   8   15   11   10   8   13   11   19   16   29   42   32   36
## 2011 2012
##   42   28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12    7    5    9    7   10    6   12    8   14   16   28   34   28   30
## 2011 2012

```

```
## 29 23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 6 5 9 7 10 6 12 8 14 13 25 32 24 30
## 2011 2012
## 24 21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 8.4, df = 16, p-value = 0.9
```



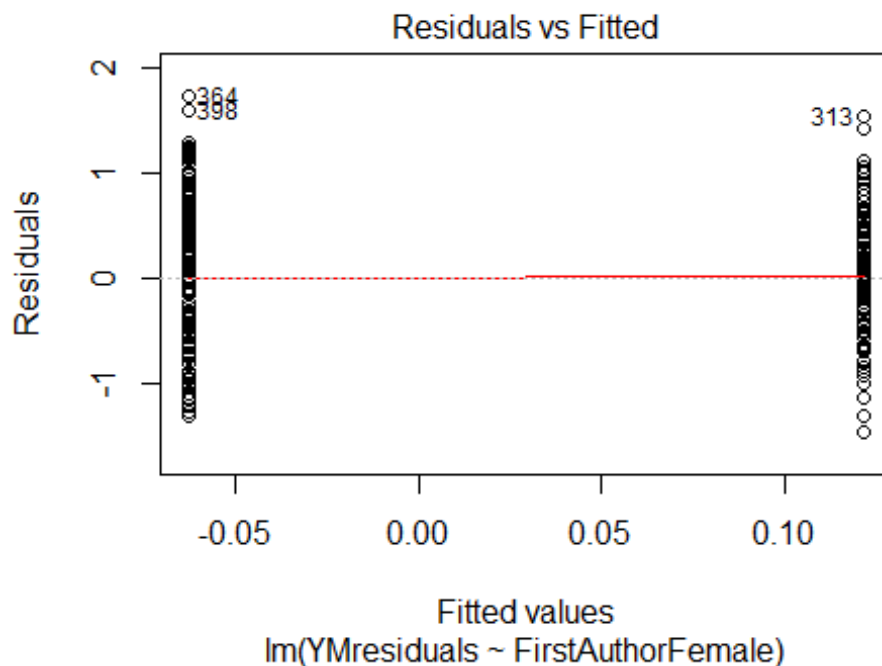
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 2.0800838230519"
## [1] "Male first author team size 2018 geometric mean: 1.62238960361098"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

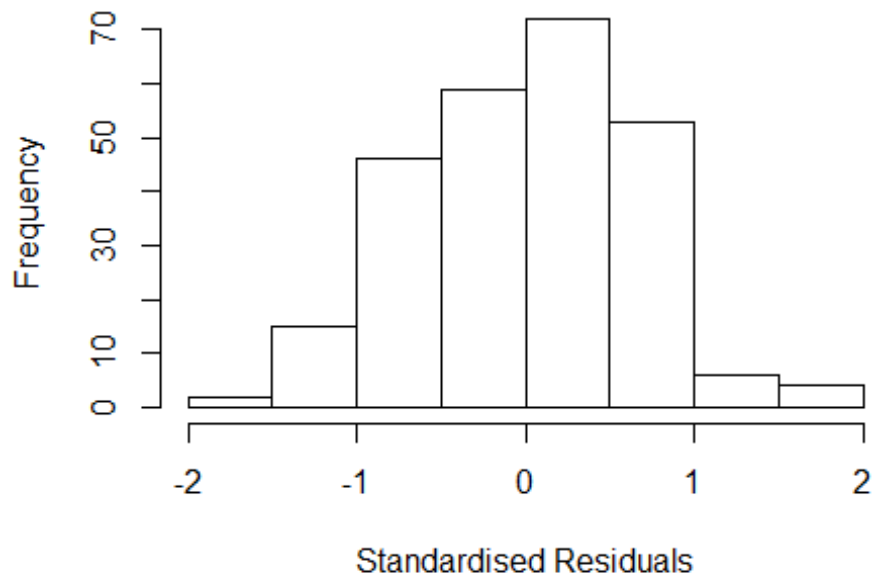
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 15, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.73205080756888"
## [1] "Male last author team size 2018 geometric mean: 1.7370729389807"

## 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 = 9, 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  5.729  1      2.394
## LastAuthorFemale  5.940  1      2.437
## UniqueAuthors    44.543  4      1.607
## Year              58.697 16      1.136
```


Residuals from first and last author and team size



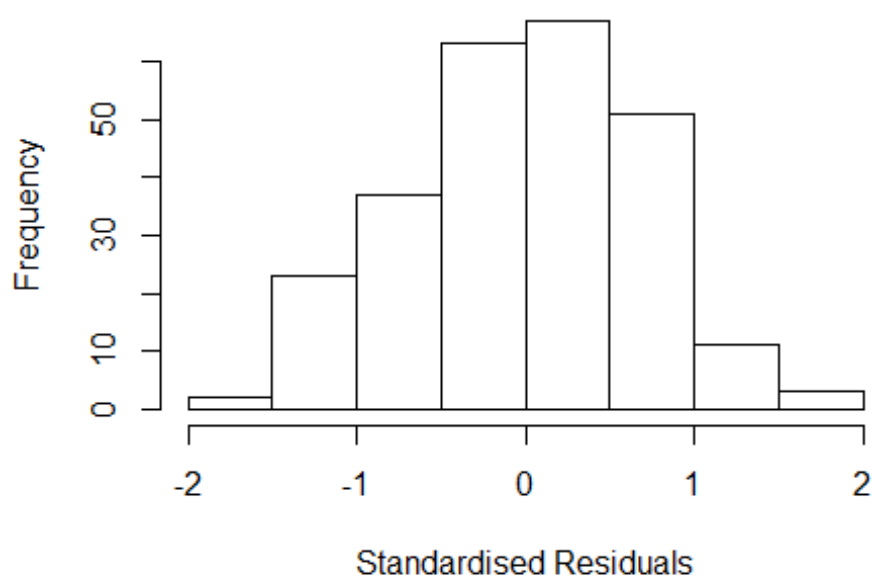
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6421 -0.4817 0.0439 0.4922 1.6577
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8158 0.1898 4.30 2.5e-05 ***
## FirstAuthorFemale1 0.0202 0.1231 0.16 0.86980
## LastAuthorFemale1 0.3589 0.1236 2.90 0.00403 **
## UniqueAuthors2 0.2897 0.0993 2.92 0.00388 **
## UniqueAuthors3 0.3816 0.1424 2.68 0.00787 **
## UniqueAuthors4 0.6245 0.1689 3.70 0.00027 ***
## UniqueAuthors5 -0.1019 0.2226 -0.46 0.64766
## Year1997 0.2235 0.3363 0.66 0.50704
## Year1998 0.8236 0.4217 1.95 0.05202 .
## Year1999 -0.0903 0.3318 -0.27 0.78577
```

```

## Year2000      0.4352      0.2757      1.58  0.11577
## Year2001      0.1232      0.2627      0.47  0.63962
## Year2002      0.2828      0.3197      0.88  0.37725
## Year2003     -0.2126      0.2672     -0.80  0.42710
## Year2004     -0.0427      0.2618     -0.16  0.87064
## Year2005     -0.0167      0.2356     -0.07  0.94348
## Year2006      0.1250      0.2668      0.47  0.63980
## Year2007     -0.1387      0.2293     -0.60  0.54594
## Year2008      0.0361      0.2244      0.16  0.87245
## Year2009      0.1777      0.2233      0.80  0.42684
## Year2010     -0.3011      0.2254     -1.34  0.18289
## Year2011      0.0968      0.2558      0.38  0.70537
## Year2012     -0.0310      0.2700     -0.11  0.90872
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.701
## Multiple R-squared:  0.185, Adjusted R-squared:  0.108
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 233 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.555  0.885  0.950  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      3.89e-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 2.239 1      1.496
## LastAuthorFemale 2.394 1      1.547
## Year      1.829 16      1.019

```

Residuals from first and last author



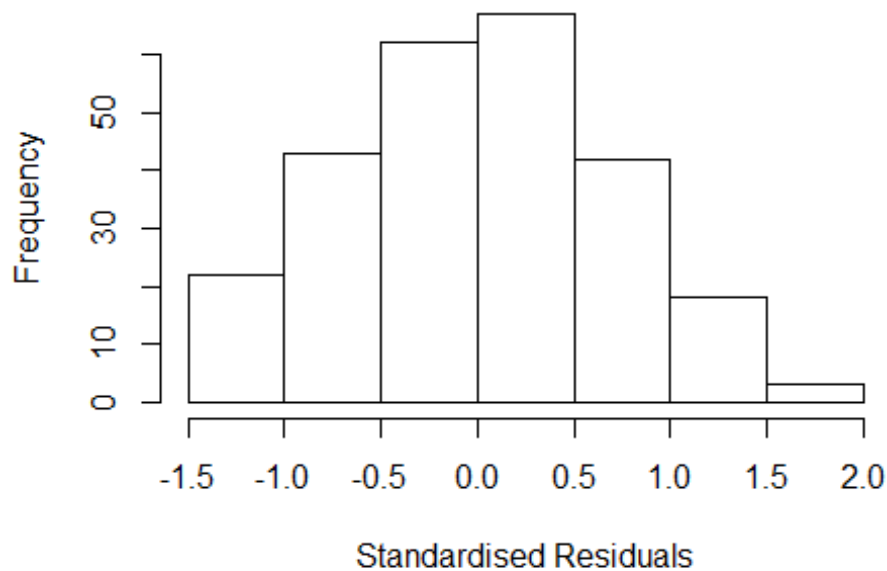
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.53706 -0.48178 0.00768 0.50769 1.79488
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.866816 0.161312 5.37 1.8e-07 ***
## FirstAuthorFemale1 0.032606 0.129158 0.25 0.8009
## LastAuthorFemale1 0.384889 0.132777 2.90 0.0041 **
## Year1997 0.209035 0.346057 0.60 0.5464
## Year1998 0.877707 0.384663 2.28 0.0234 *
## Year1999 -0.007316 0.305481 -0.02 0.9809
## Year2000 0.500085 0.238750 2.09 0.0373 *
## Year2001 0.215860 0.269415 0.80 0.4238
## Year2002 0.330741 0.348563 0.95 0.3436
## Year2003 -0.146616 0.275469 -0.53 0.5951
## Year2004 0.000882 0.241916 0.00 0.9971
## Year2005 0.135010 0.224993 0.60 0.5490
```

```

## Year2006          0.206322    0.225432    0.92    0.3610
## Year2007          -0.063492    0.216985   -0.29    0.7701
## Year2008          0.150119    0.198389    0.76    0.4500
## Year2009          0.285355    0.190926    1.49    0.1363
## Year2010          -0.171304    0.199118   -0.86    0.3905
## Year2011          0.217332    0.249776    0.87    0.3851
## Year2012          0.120084    0.248370    0.48    0.6292
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.72
## Multiple R-squared:  0.13,    Adjusted R-squared:  0.0645
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 237 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.514  0.889  0.949  0.919  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      3.89e-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 1.341 1      1.158
## Year              1.341 16      1.009

```

Residuals from first author



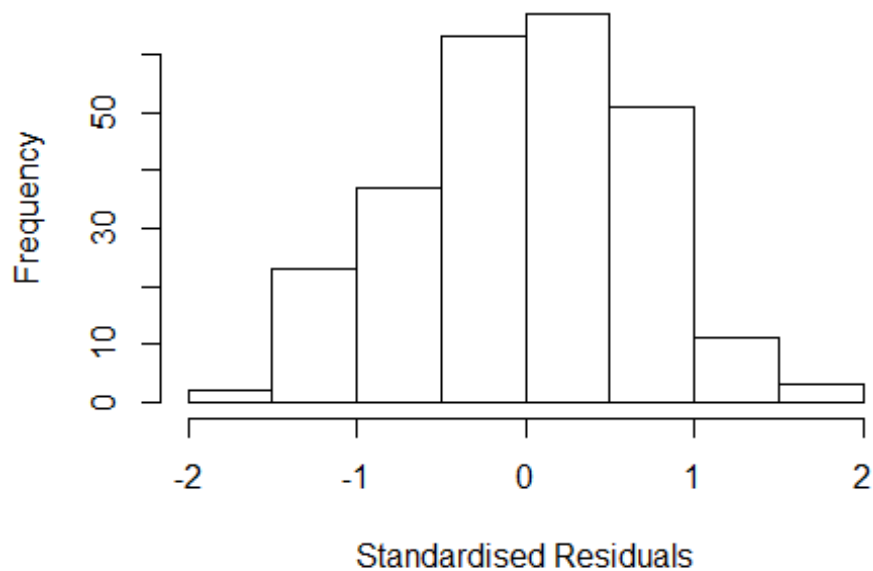
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.40583 -0.50607 0.00143 0.48841 1.82226
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9526 0.1660 5.74 2.9e-08 ***
## FirstAuthorFemale1 0.2321 0.1010 2.30 0.0225 *
## Year1997 0.1273 0.3332 0.38 0.7027
## Year1998 0.8582 0.3292 2.61 0.0097 **
## Year1999 -0.0625 0.3203 -0.20 0.8454
## Year2000 0.5592 0.2517 2.22 0.0273 *
## Year2001 0.2282 0.2720 0.84 0.4024
## Year2002 0.3830 0.3526 1.09 0.2784
## Year2003 -0.0962 0.2890 -0.33 0.7395
## Year2004 0.0134 0.2615 0.05 0.9591
## Year2005 0.0482 0.2183 0.22 0.8255
## Year2006 0.2576 0.2405 1.07 0.2852
```

```

## Year2007          -0.0619      0.2153   -0.29   0.7738
## Year2008          0.1004      0.2104    0.48   0.6336
## Year2009          0.2010      0.1911    1.05   0.2939
## Year2010         -0.1697      0.2075   -0.82   0.4143
## Year2011          0.2212      0.2600    0.85   0.3958
## Year2012          0.0663      0.2546    0.26   0.7949
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.769
## Multiple R-squared:  0.0856, Adjusted R-squared:  0.0205
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 237 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.554  0.889   0.948   0.926   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      3.89e-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 1.413 1          1.189
## Year            1.413 16          1.011

```

Residuals from last author



```
## [1] "List of 0 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
## -1.5606 -0.4738 0.0137 0.5097 1.8231
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.87500 0.15991 5.47 1.1e-07 ***
## LastAuthorFemale1 0.39972 0.10305 3.88 0.00014 ***
## Year1997 0.20916 0.34871 0.60 0.54920
## Year1998 0.86613 0.38467 2.25 0.02525 *
## Year1999 -0.01009 0.30266 -0.03 0.97344
## Year2000 0.49018 0.23358 2.10 0.03691 *
## Year2001 0.21639 0.26875 0.81 0.42153
## Year2002 0.32516 0.34348 0.95 0.34476
## Year2003 -0.14375 0.27356 -0.53 0.59973
## Year2004 0.00201 0.24299 0.01 0.99342
## Year2005 0.13435 0.22582 0.59 0.55244
## Year2006 0.19629 0.22089 0.89 0.37510
```

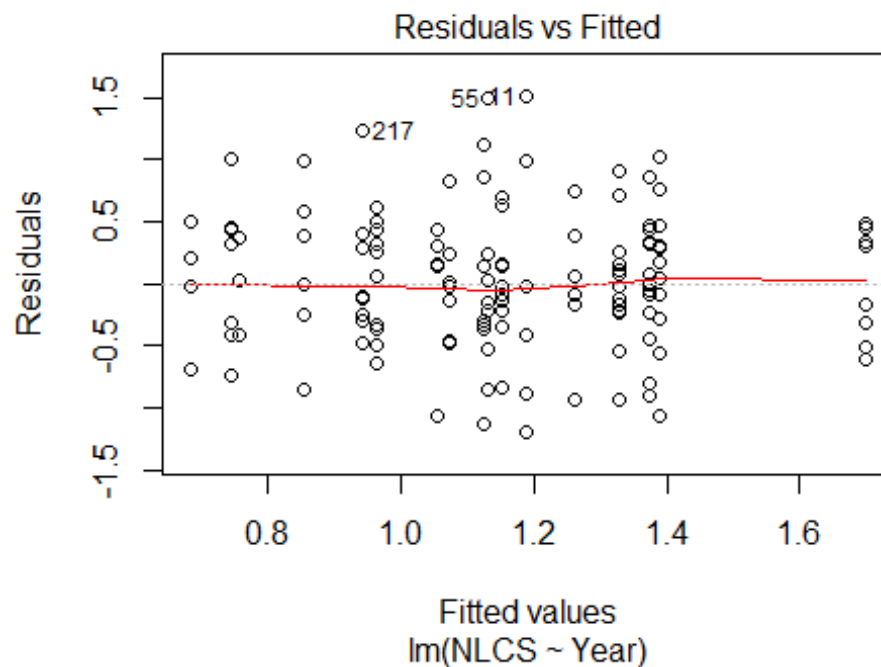
```

## Year2007          -0.06881      0.21614      -0.32   0.75048
## Year2008           0.14480      0.19664       0.74   0.46222
## Year2009           0.28584      0.19024       1.50   0.13428
## Year2010          -0.17508      0.19769      -0.89   0.37672
## Year2011           0.21860      0.24854       0.88   0.38000
## Year2012           0.12535      0.24709       0.51   0.61239
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.734
## Multiple R-squared:  0.129, Adjusted R-squared:  0.0672
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 234 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.517  0.891  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      3.89e-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: 257"
## [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
##    7    3    6    8    8   10   11   14    8    9   13   16    8   12   12
## 2011 2012
##   17   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    3    6    6    5    8    7   11    4    8   12   13    7   10    7
## 2011 2012

```



```
## 13 11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 3 5 6 5 7 7 10 3 7 12 13 5 10 7
## 2011 2012
## 12 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```



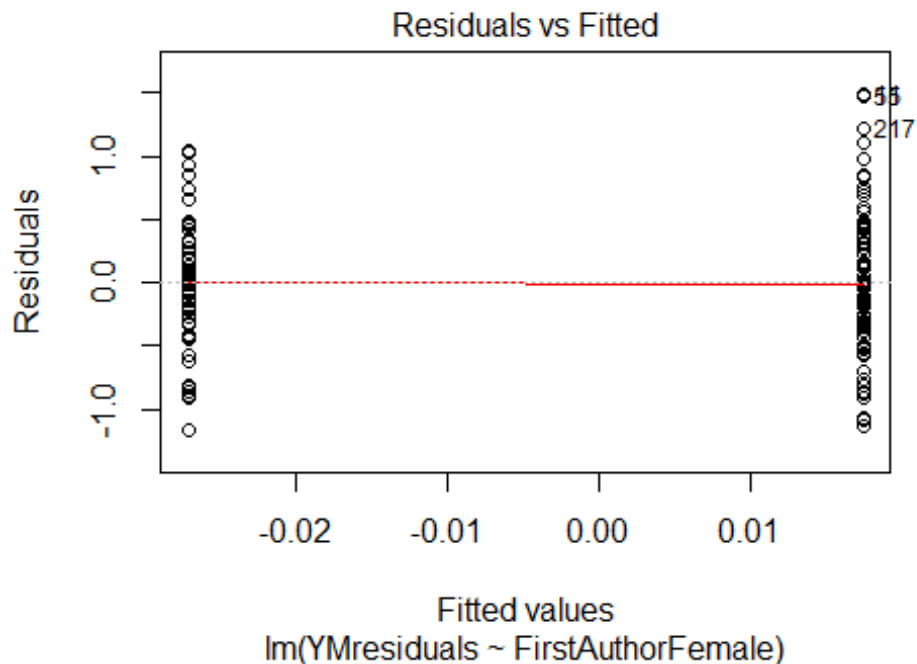
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.67, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 2.91295063024394"
## [1] "Male first author team size 2018 geometric mean: 2.6207413942089"

## 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 = 12, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.71080601082953"
## [1] "Male last author team size 2018 geometric mean: 2.74945927399721"

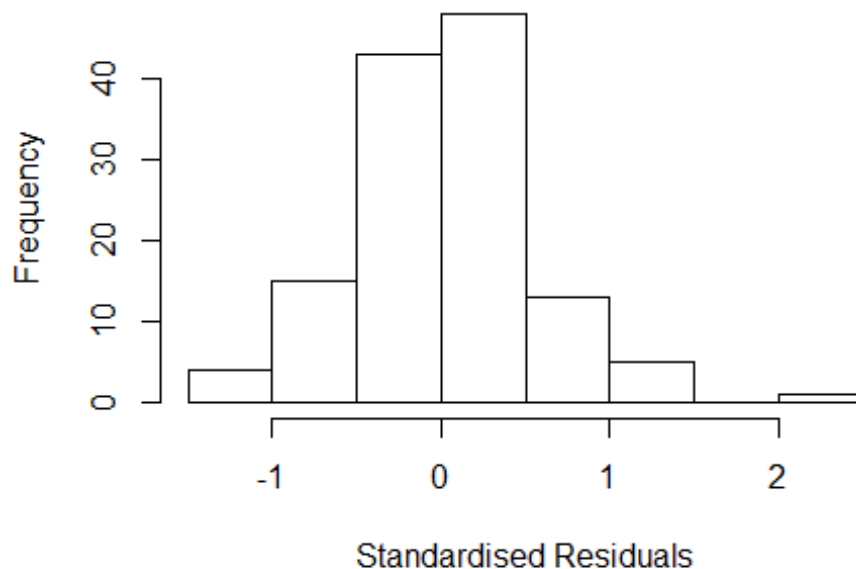
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8.5, 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	3.243	1	1.801
LastAuthorFemale	8.377	1	2.894
UniqueAuthors	19.219	4	1.447
Year	66.693	16	1.140

Residuals from first and last author and team size



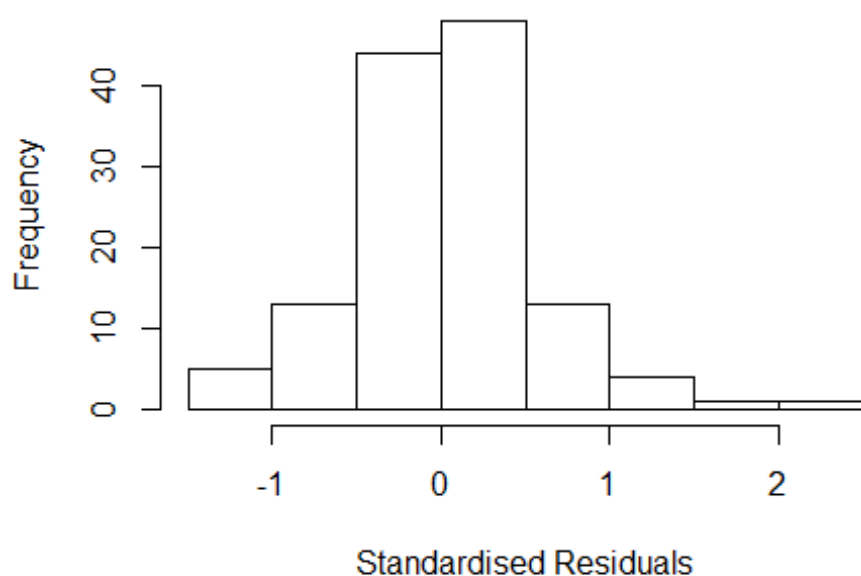
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2963 -0.3092 0.0333 0.2856 2.0725
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.90568 0.46323 1.96 0.053 .
## FirstAuthorFemale1 -0.15759 0.12930 -1.22 0.226
## LastAuthorFemale1 0.08925 0.18816 0.47 0.636
## UniqueAuthors2 0.12906 0.13425 0.96 0.339
## UniqueAuthors3 0.21491 0.20728 1.04 0.302
## UniqueAuthors4 -0.35600 1.70906 -0.21 0.835
## UniqueAuthors5 0.28135 0.15394 1.83 0.070 .
## Year1997 -0.14639 0.50569 -0.29 0.773
## Year1998 -0.41328 0.54750 -0.75 0.452
## Year1999 0.39049 0.49488 0.79 0.432
```

```

## Year2000      0.15555      0.47454      0.33      0.744
## Year2001     -0.18677      0.53132     -0.35      0.726
## Year2002      0.00255      0.45907      0.01      0.996
## Year2003     -0.02136      0.43442     -0.05      0.961
## Year2004     -0.14251      0.41388     -0.34      0.731
## Year2005      0.82689      0.44465      1.86      0.066 .
## Year2006      0.49569      0.46141      1.07      0.285
## Year2007      0.36872      0.40873      0.90      0.369
## Year2008      0.10227      0.54547      0.19      0.852
## Year2009      0.18489      0.43400      0.43      0.671
## Year2010     -0.01707      0.41190     -0.04      0.967
## Year2011      0.44163      0.42844      1.03      0.305
## Year2012     -0.04376      0.41868     -0.10      0.917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.271, Adjusted R-squared:  0.12
## Convergence in 35 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 116 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.007  0.816  0.957  0.878  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      7.75e-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 2.832 1      1.683
## LastAuthorFemale  4.671 1      2.161
## Year              6.647 16      1.061

```

Residuals from first and last author



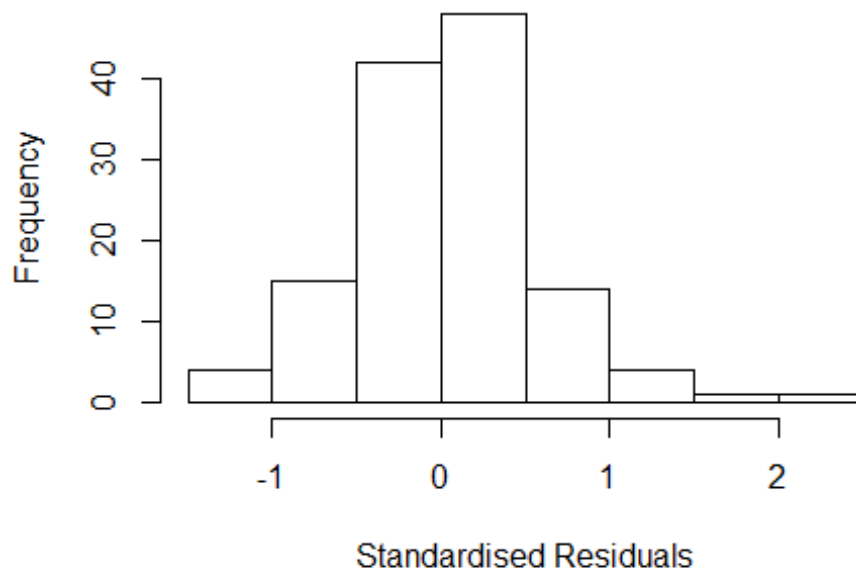
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3511 -0.2860 0.0251 0.3034 2.1195
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09291 0.38666 2.83 0.0056 **
## FirstAuthorFemale1 -0.16373 0.13129 -1.25 0.2150
## LastAuthorFemale1 0.12782 0.15105 0.85 0.3993
## Year1997 -0.33383 0.43441 -0.77 0.4439
## Year1998 -0.51845 0.46341 -1.12 0.2657
## Year1999 0.21811 0.41099 0.53 0.5967
## Year2000 0.13039 0.45246 0.29 0.7737
## Year2001 -0.29731 0.52244 -0.57 0.5705
## Year2002 -0.14987 0.40521 -0.37 0.7122
## Year2003 -0.13129 0.38613 -0.34 0.7345
## Year2004 -0.21275 0.37278 -0.57 0.5694
## Year2005 0.66824 0.39463 1.69 0.0932 .
```

```

## Year2006          0.35187    0.43427    0.81    0.4195
## Year2007          0.28988    0.38044    0.76    0.4477
## Year2008         -0.00731    0.55228   -0.01    0.9895
## Year2009          0.11188    0.39379    0.28    0.7769
## Year2010         -0.09072    0.39581   -0.23    0.8191
## Year2011          0.30017    0.37260    0.81    0.4222
## Year2012         -0.21598    0.36493   -0.59    0.5552
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.482
## Multiple R-squared:  0.248, Adjusted R-squared:  0.125
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 115 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0145 0.8560 0.9570 0.8830 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      7.75e-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 1.84 1      1.357
## Year              1.84 16      1.019

```

Residuals from first author



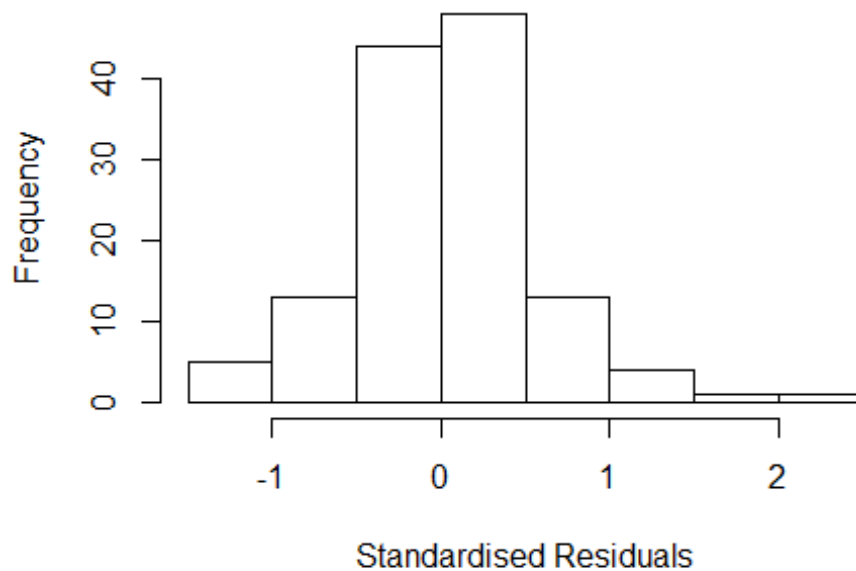
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.19044 -0.33665 0.00895 0.35914 2.06135
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1330 0.3533 3.21 0.0018 **
## FirstAuthorFemale1 -0.1049 0.1117 -0.94 0.3500
## Year1997 -0.3743 0.4032 -0.93 0.3553
## Year1998 -0.5004 0.4665 -1.07 0.2858
## Year1999 0.2006 0.3944 0.51 0.6121
## Year2000 0.0574 0.4177 0.14 0.8909
## Year2001 -0.3348 0.4748 -0.71 0.4822
## Year2002 -0.1251 0.3979 -0.31 0.7538
## Year2003 -0.1459 0.3694 -0.39 0.6937
## Year2004 -0.1868 0.3637 -0.51 0.6086
## Year2005 0.6845 0.3803 1.80 0.0745 .
## Year2006 0.3193 0.4067 0.79 0.4341
```

```

## Year2007          0.2964      0.3652      0.81      0.4188
## Year2008         -0.0551      0.4962     -0.11      0.9117
## Year2009          0.0687      0.3624      0.19      0.8499
## Year2010         -0.0417      0.3755     -0.11      0.9119
## Year2011          0.3020      0.3540      0.85      0.3955
## Year2012         -0.2038      0.3533     -0.58      0.5653
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.219, Adjusted R-squared:  0.0999
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 114 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0824 0.8650 0.9550 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      7.75e-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 3.477 1          1.865
## Year              3.477 16          1.040

```


Residuals from last author



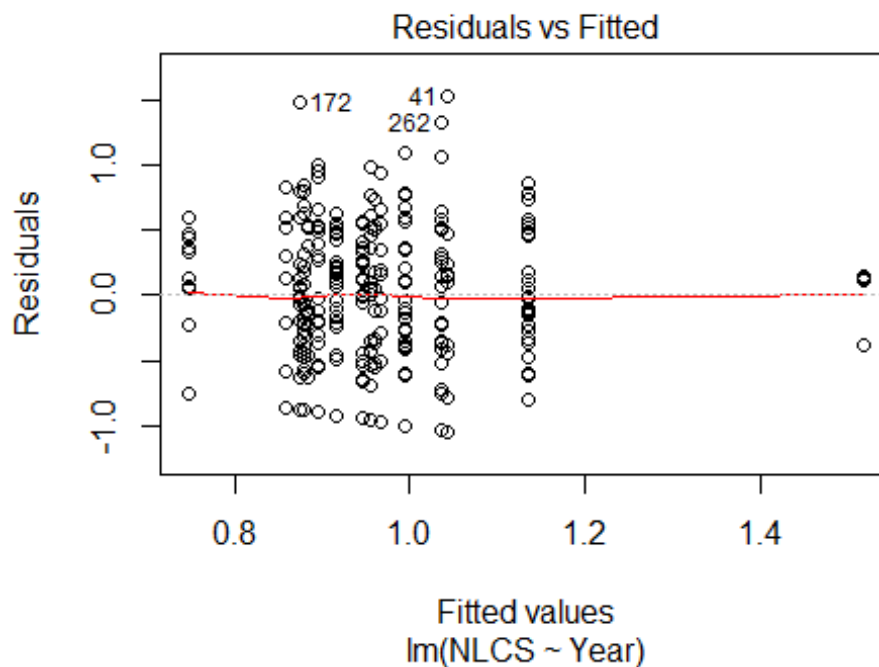
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.179 -0.296 0.020 0.327 2.125
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0575 0.3693 2.86 0.005 **
## LastAuthorFemale1 0.0371 0.1332 0.28 0.781
## Year1997 -0.2985 0.4186 -0.71 0.477
## Year1998 -0.4883 0.4700 -1.04 0.301
## Year1999 0.2124 0.4114 0.52 0.607
## Year2000 0.0844 0.4324 0.20 0.846
## Year2001 -0.2616 0.4874 -0.54 0.593
## Year2002 -0.1028 0.3944 -0.26 0.795
## Year2003 -0.1415 0.3764 -0.38 0.708
## Year2004 -0.1729 0.3578 -0.48 0.630
## Year2005 0.6746 0.3846 1.75 0.082 .
## Year2006 0.3632 0.4136 0.88 0.382
```

```

## Year2007          0.3105      0.3706      0.84      0.404
## Year2008          0.0250      0.5454      0.05      0.963
## Year2009          0.0900      0.3898      0.23      0.818
## Year2010         -0.0373      0.3672     -0.10      0.919
## Year2011          0.2923      0.3616      0.81      0.421
## Year2012         -0.2464      0.3550     -0.69      0.489
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.226, Adjusted R-squared:  0.108
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 110 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0288 0.8430 0.9500 0.8830 0.9810 0.9980
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      7.75e-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: 129"
## [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
##   13   5   9  12  13  14  11  11  10  21  21  25  32  26  37
## 2011 2012
##   41  33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12   4   9  10  12  12  10  11  10  18  17  24  26  22  30
## 2011 2012

```

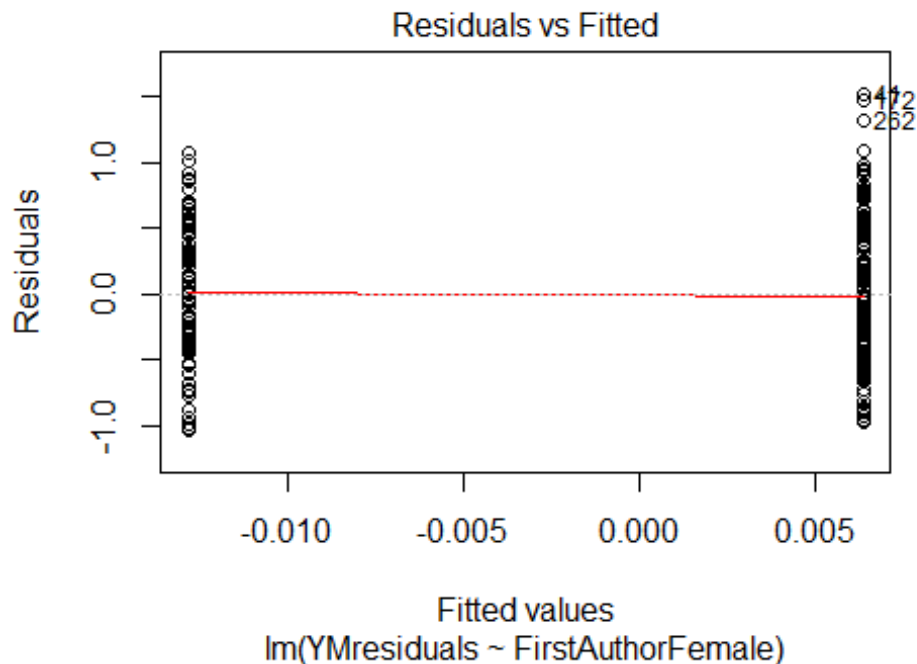
```
## 32 28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 4 9 10 12 12 9 11 10 18 15 23 24 22 25
## 2011 2012
## 29 26
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.073, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 1.5874010519682"
## [1] "Male first author team size 2018 geometric mean: 2.00730966250204"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

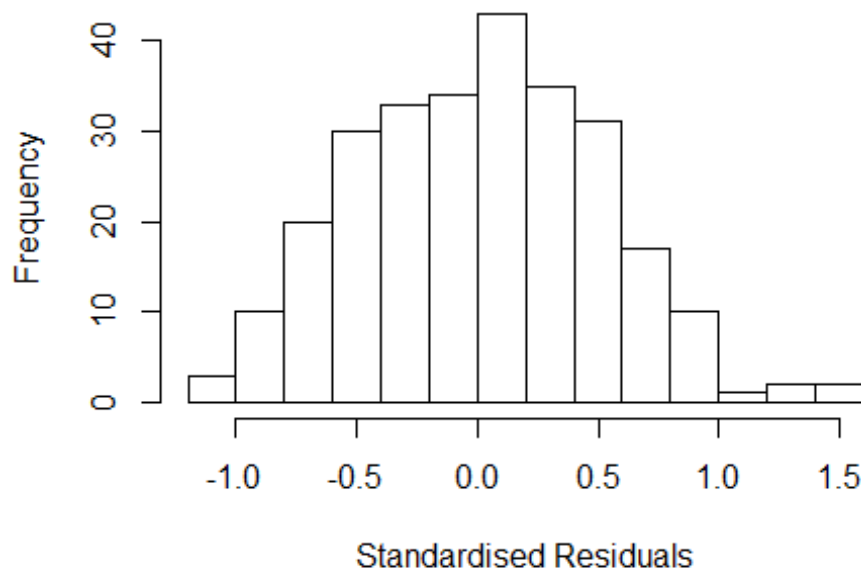
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 58, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.76172958987204"
## [1] "Male last author team size 2018 geometric mean: 1.91421605846746"

## 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 = 78, 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 1.651  1      1.285
## LastAuthorFemale  1.367  1      1.169
## UniqueAuthors    2.377  3      1.155
## Year              3.792 16      1.043
```

Residuals from first and last author and team size



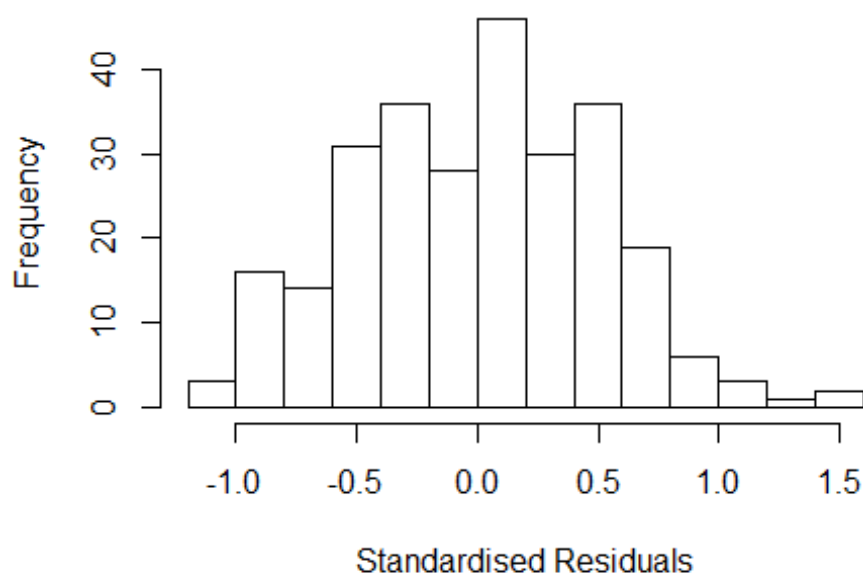
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0978 -0.3697 0.0234 0.3745 1.4291
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.84797 0.18145 4.67 4.9e-06 ***
## FirstAuthorFemale1 -0.03906 0.08118 -0.48 0.6308
## LastAuthorFemale1 0.08633 0.08067 1.07 0.2856
## UniqueAuthors2 0.21868 0.07717 2.83 0.0050 **
## UniqueAuthors3 0.23637 0.10000 2.36 0.0189 *
## UniqueAuthors4 0.81651 0.10715 7.62 5.3e-13 ***
## Year1997 0.58064 0.19893 2.92 0.0038 **
## Year1998 0.04312 0.24513 0.18 0.8605
## Year1999 0.07025 0.27396 0.26 0.7978
## Year2000 -0.15933 0.23247 -0.69 0.4938
```

```

## Year2001      0.00445    0.21227    0.02    0.9833
## Year2002      0.00341    0.25675    0.01    0.9894
## Year2003     -0.09875    0.21410   -0.46    0.6450
## Year2004     -0.10407    0.21619   -0.48    0.6307
## Year2005     -0.10360    0.22208   -0.47    0.6413
## Year2006     -0.14105    0.22302   -0.63    0.5277
## Year2007     -0.14885    0.19990   -0.74    0.4572
## Year2008     -0.10615    0.19274   -0.55    0.5823
## Year2009     -0.04837    0.21930   -0.22    0.8256
## Year2010      0.11584    0.19640    0.59    0.5559
## Year2011     -0.15162    0.21590   -0.70    0.4832
## Year2012      0.02038    0.21773    0.09    0.9255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.531
## Multiple R-squared:  0.104, Adjusted R-squared:  0.0281
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 242 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.449  0.875   0.946   0.914   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      3.69e-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 1.447 1      1.203
## LastAuthorFemale  1.424 1      1.193
## Year              1.617 16      1.015

```

Residuals from first and last author



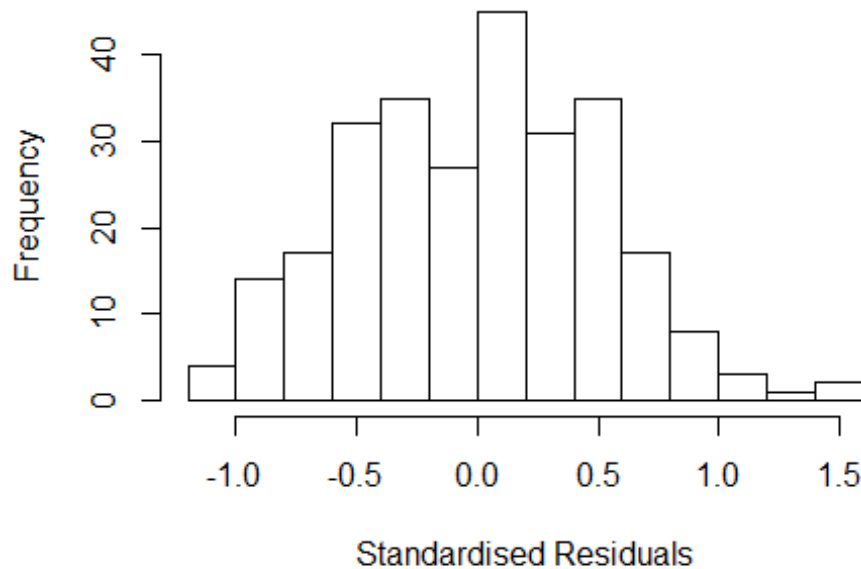
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0495 -0.3920 0.0394 0.3857 1.5958
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98851 0.18678 5.29 2.6e-07 ***
## FirstAuthorFemale1 -0.03014 0.08287 -0.36 0.716
## LastAuthorFemale1 0.06784 0.08445 0.80 0.423
## Year1997 0.54821 0.22004 2.49 0.013 *
## Year1998 -0.03466 0.25696 -0.13 0.893
## Year1999 -0.01829 0.28277 -0.06 0.948
## Year2000 -0.22490 0.23996 -0.94 0.350
## Year2001 -0.02449 0.21922 -0.11 0.911
## Year2002 -0.00643 0.26478 -0.02 0.981
## Year2003 -0.05437 0.23301 -0.23 0.816
## Year2004 -0.12814 0.22938 -0.56 0.577
## Year2005 -0.04740 0.22987 -0.21 0.837
```

```

## Year2006      -0.14525    0.23304   -0.62    0.534
## Year2007      -0.13042    0.21797   -0.60    0.550
## Year2008      -0.04497    0.20921   -0.21    0.830
## Year2009       0.02332    0.23271    0.10    0.920
## Year2010       0.10522    0.20842    0.50    0.614
## Year2011      -0.15889    0.22863   -0.69    0.488
## Year2012       0.01556    0.22161    0.07    0.944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.0439, Adjusted R-squared:  -0.0244
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 253 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.378  0.893   0.947   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      3.69e-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 1.267 1      1.125
## Year              1.267 16      1.007

```


Residuals from first author



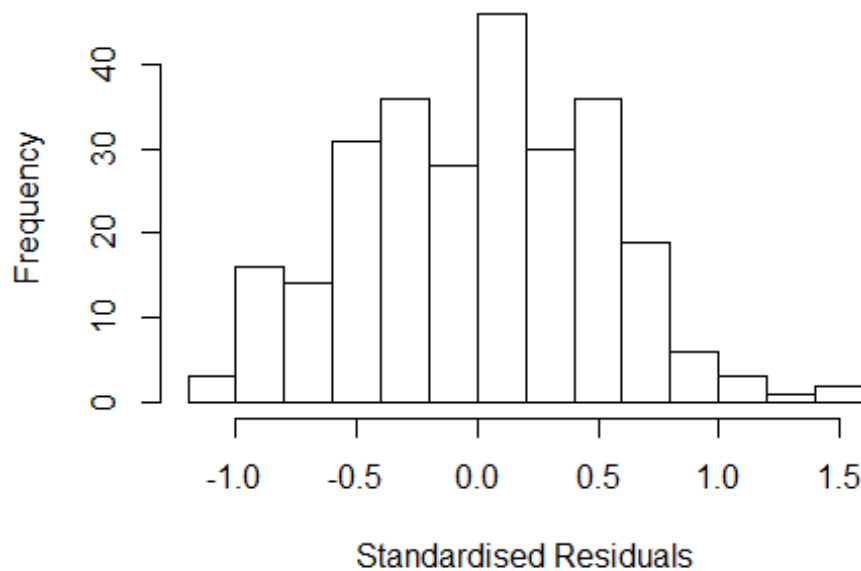
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0200 -0.3883 0.0334 0.3793 1.5879
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98629 0.18685 5.28 2.8e-07 ***
## FirstAuthorFemale1 -0.00304 0.07758 -0.04 0.969
## Year1997 0.53645 0.21787 2.46 0.014 *
## Year1998 -0.03244 0.25704 -0.13 0.900
## Year1999 -0.00814 0.27978 -0.03 0.977
## Year2000 -0.21670 0.23934 -0.91 0.366
## Year2001 -0.01306 0.21859 -0.06 0.952
## Year2002 -0.00800 0.26445 -0.03 0.976
## Year2003 -0.03538 0.23166 -0.15 0.879
## Year2004 -0.10362 0.22921 -0.45 0.652
## Year2005 -0.03124 0.23141 -0.13 0.893
## Year2006 -0.12476 0.23340 -0.53 0.593
```

```

## Year2007          -0.10938    0.21701   -0.50    0.615
## Year2008          -0.02325    0.20398   -0.11    0.909
## Year2009           0.03679    0.23501    0.16    0.876
## Year2010           0.12056    0.20733    0.58    0.561
## Year2011          -0.14872    0.22746   -0.65    0.514
## Year2012           0.02785    0.22093    0.13    0.900
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared:  0.0414, Adjusted R-squared:  -0.0231
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 251 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.383  0.890  0.945  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      3.69e-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 1.268 1      1.126
## Year              1.268 16      1.007

```

Residuals from last author



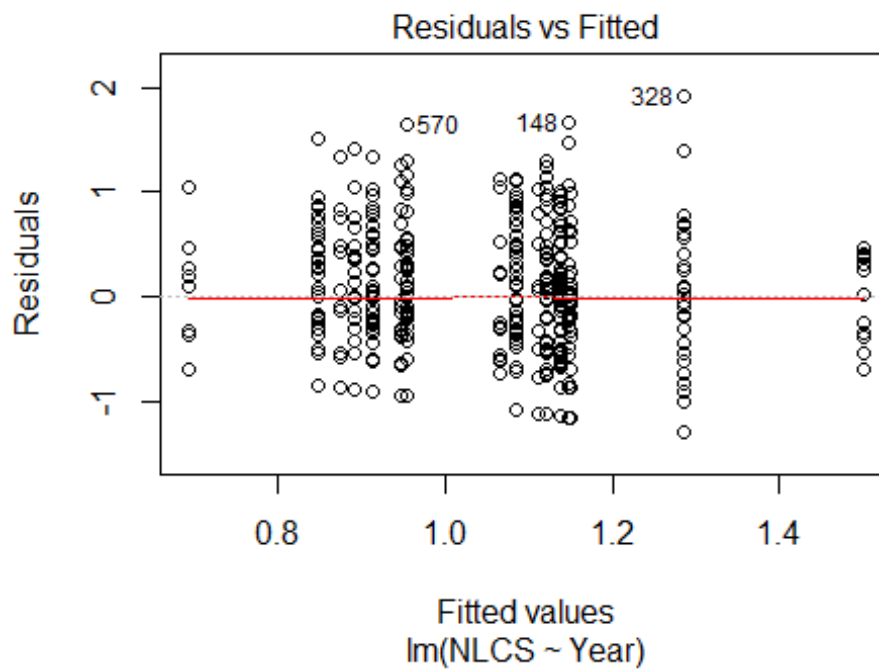
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0603 -0.3838 0.0424 0.3765 1.5996
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98083 0.18226 5.38 1.7e-07 ***
## LastAuthorFemale1 0.05496 0.07959 0.69 0.491
## Year1997 0.54026 0.21384 2.53 0.012 *
## Year1998 -0.02716 0.25332 -0.11 0.915
## Year1999 -0.01441 0.28322 -0.05 0.959
## Year2000 -0.22020 0.23693 -0.93 0.354
## Year2001 -0.02911 0.21596 -0.13 0.893
## Year2002 -0.00344 0.26203 -0.01 0.990
## Year2003 -0.04539 0.22824 -0.20 0.843
## Year2004 -0.12561 0.22685 -0.55 0.580
## Year2005 -0.04442 0.22818 -0.19 0.846
## Year2006 -0.14312 0.23033 -0.62 0.535
```

```

## Year2007          -0.13123      0.21483    -0.61      0.542
## Year2008          -0.03924      0.20475    -0.19      0.848
## Year2009           0.02453      0.23019      0.11      0.915
## Year2010           0.10638      0.20568      0.52      0.605
## Year2011          -0.15828      0.22570    -0.70      0.484
## Year2012           0.01499      0.21920      0.07      0.946
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.554
## Multiple R-squared:  0.0433, Adjusted R-squared:  -0.021
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 250 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.384  0.895  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      3.69e-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: 271"
## [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
##   12   17   15   15   27   17   25   27   38   23   34   54   48   55   61
## 2011 2012
##   63   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9   12   11   13   21   14   21   21   28   20   29   49   32   46   51
## 2011 2012

```

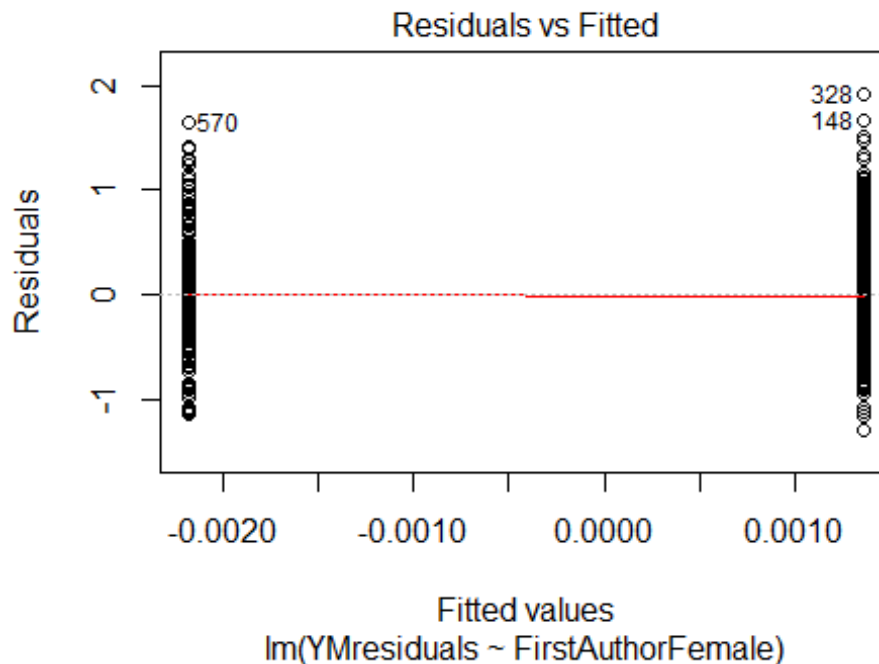
```
## 53 42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 11 11 13 21 10 20 18 27 18 23 44 28 44 43
## 2011 2012
## 45 34
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 12, df = 16, p-value = 0.7
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.7, df = 1, p-value = 0.4
## [1] "Female first author team size 2018 geometric mean: 2.31201044739033"
## [1] "Male first author team size 2018 geometric mean: 1.8364704589087"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

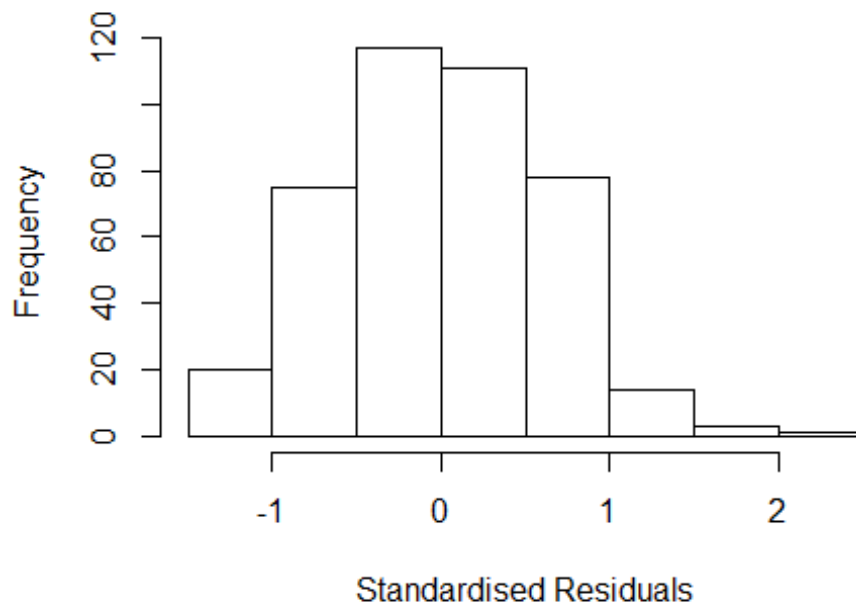
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.27963274006424"
## [1] "Male last author team size 2018 geometric mean: 1.87220300352478"

## 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 = 0.5
## 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 2.283  1      1.511
## LastAuthorFemale  2.157  1      1.469
## UniqueAuthors    1.484  4      1.051
## Year              1.733 16      1.017
```

Residuals from first and last author and team size



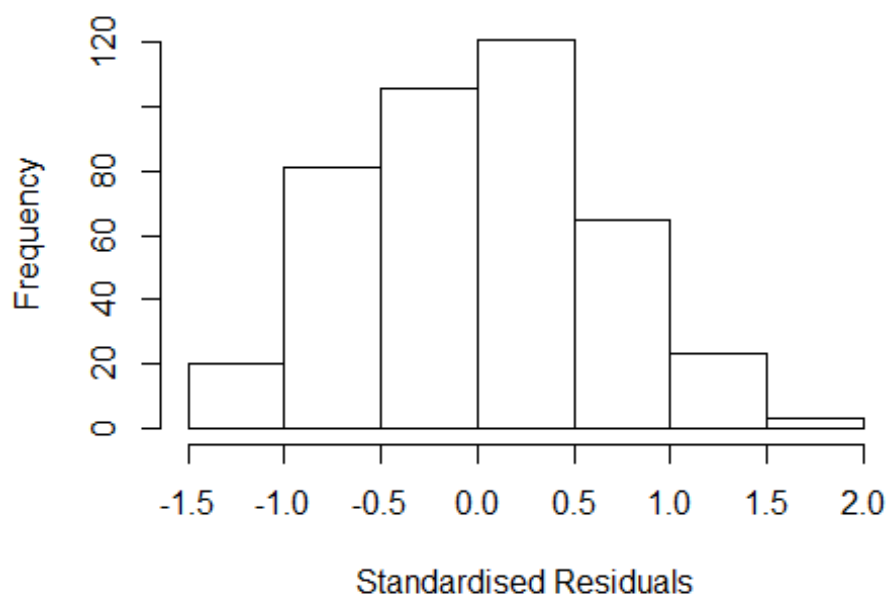
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3002 -0.4258 -0.0073 0.4373 2.0079
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6631 0.1824 3.63 0.00031 ***
## FirstAuthorFemale1 -0.0764 0.0929 -0.82 0.41123
## LastAuthorFemale1 0.0407 0.0908 0.45 0.65396
## UniqueAuthors2 0.2637 0.0715 3.69 0.00026 ***
## UniqueAuthors3 0.1082 0.1339 0.81 0.41976
## UniqueAuthors4 0.5579 0.1999 2.79 0.00553 **
## UniqueAuthors5 -0.0718 0.5076 -0.14 0.88753
## Year1997 0.7248 0.2210 3.28 0.00113 **
## Year1998 0.4177 0.2727 1.53 0.12640
## Year1999 0.3115 0.2437 1.28 0.20180
```

```

## Year2000          0.4099      0.2316      1.77  0.07751 .
## Year2001         -0.1042      0.2451     -0.43  0.67079
## Year2002          0.3126      0.2318      1.35  0.17823
## Year2003          0.1723      0.2435      0.71  0.47957
## Year2004          0.1137      0.2123      0.54  0.59257
## Year2005          0.4204      0.2175      1.93  0.05402 .
## Year2006          0.5289      0.2339      2.26  0.02429 *
## Year2007          0.1353      0.2021      0.67  0.50366
## Year2008          0.4262      0.2056      2.07  0.03882 *
## Year2009          0.0864      0.2091      0.41  0.67971
## Year2010          0.1267      0.2035      0.62  0.53390
## Year2011          0.3427      0.2114      1.62  0.10577
## Year2012          0.3028      0.2191      1.38  0.16785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.672
## Multiple R-squared:  0.107, Adjusted R-squared:  0.0569
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 388 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.352  0.880  0.953  0.923  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      2.39e-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 2.195  1      1.481
## LastAuthorFemale  2.200  1      1.483
## Year              1.335 16      1.009

```


Residuals from first and last author



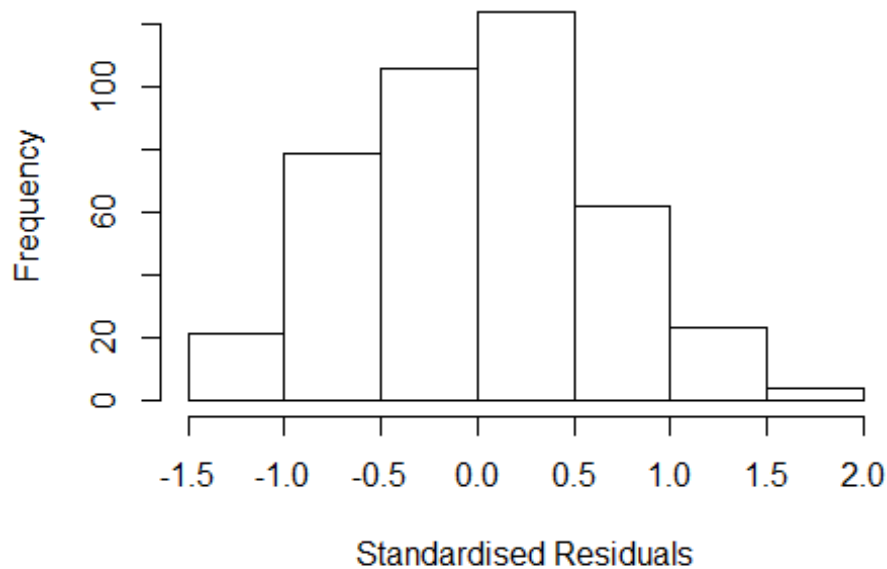
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.28570 -0.46177 0.00649 0.43548 1.91430
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6671 0.1954 3.41 0.00071 ***
## FirstAuthorFemale1 -0.0219 0.0939 -0.23 0.81550
## LastAuthorFemale1 0.0457 0.0941 0.49 0.62766
## Year1997 0.8643 0.2338 3.70 0.00025 ***
## Year1998 0.4365 0.2768 1.58 0.11564
## Year1999 0.3612 0.2586 1.40 0.16322
## Year2000 0.4713 0.2423 1.94 0.05248 .
## Year2001 -0.0739 0.2681 -0.28 0.78283
## Year2002 0.3594 0.2479 1.45 0.14794
## Year2003 0.2444 0.2509 0.97 0.33050
## Year2004 0.2111 0.2249 0.94 0.34841
## Year2005 0.4902 0.2313 2.12 0.03472 *
```

```

## Year2006          0.6186      0.2404      2.57  0.01043 *
## Year2007          0.2094      0.2133      0.98  0.32682
## Year2008          0.4945      0.2112      2.34  0.01970 *
## Year2009          0.1694      0.2167      0.78  0.43483
## Year2010          0.2400      0.2174      1.10  0.27027
## Year2011          0.4217      0.2234      1.89  0.05977 .
## Year2012          0.4169      0.2315      1.80  0.07250 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.677
## Multiple R-squared:  0.0691, Adjusted R-squared:  0.0272
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 385 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.404  0.866  0.951  0.920  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      2.39e-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 1.126 1      1.061
## Year              1.126 16      1.004

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.29179 -0.46918 0.00998 0.44322 1.90821
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.67512 0.19412 3.48 0.00056 ***
## FirstAuthorFemale1 0.00678 0.06743 0.10 0.91996
## Year1997 0.85396 0.23076 3.70 0.00025 ***
## Year1998 0.43991 0.27564 1.60 0.11129
## Year1999 0.35670 0.25882 1.38 0.16891
## Year2000 0.46221 0.24137 1.91 0.05621 .
## Year2001 -0.07337 0.26870 -0.27 0.78494
## Year2002 0.35146 0.24632 1.43 0.15440
## Year2003 0.24235 0.25111 0.97 0.33508
## Year2004 0.21125 0.22417 0.94 0.34658
## Year2005 0.49267 0.23180 2.13 0.03416 *
## Year2006 0.61668 0.23956 2.57 0.01040 *
```

```

## Year2007          0.20912      0.21296      0.98  0.32671
## Year2008          0.49783      0.21070      2.36  0.01862 *
## Year2009          0.16532      0.21556      0.77  0.44359
## Year2010          0.23530      0.21707      1.08  0.27903
## Year2011          0.42025      0.22287      1.89  0.06007 .
## Year2012          0.41229      0.22989      1.79  0.07365 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.675
## Multiple R-squared:  0.0684, Adjusted R-squared:  0.0289
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 389 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.404  0.864  0.951  0.920  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      2.39e-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 1.129 1      1.062
## 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.28015 -0.45823  0.00721  0.43699  1.91985

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6632    0.1949    3.40 0.00073 ***
## LastAuthorFemale1 0.0317    0.0677    0.47 0.63957
## Year1997        0.8626    0.2334    3.70 0.00025 ***
## Year1998        0.4390    0.2769    1.59 0.11368
## Year1999        0.3607    0.2600    1.39 0.16611
## Year2000        0.4717    0.2431    1.94 0.05305 .
## Year2001       -0.0702    0.2678   -0.26 0.79344
## Year2002        0.3596    0.2483    1.45 0.14822
## Year2003        0.2475    0.2499    0.99 0.32242
## Year2004        0.2143    0.2239    0.96 0.33920
## Year2005        0.4922    0.2316    2.13 0.03417 *
## Year2006        0.6169    0.2402    2.57 0.01059 *
## Year2007        0.2115    0.2132    0.99 0.32178
## Year2008        0.4968    0.2108    2.36 0.01892 *
## Year2009        0.1677    0.2168    0.77 0.43970
## Year2010        0.2390    0.2175    1.10 0.27256
## Year2011        0.4214    0.2234    1.89 0.06004 .
## Year2012        0.4152    0.2316    1.79 0.07374 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.677
## Multiple R-squared:  0.0689, Adjusted R-squared:  0.0294
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 385 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.401  0.866  0.951  0.920  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      2.39e-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: 419"
## [1] ""

```

```

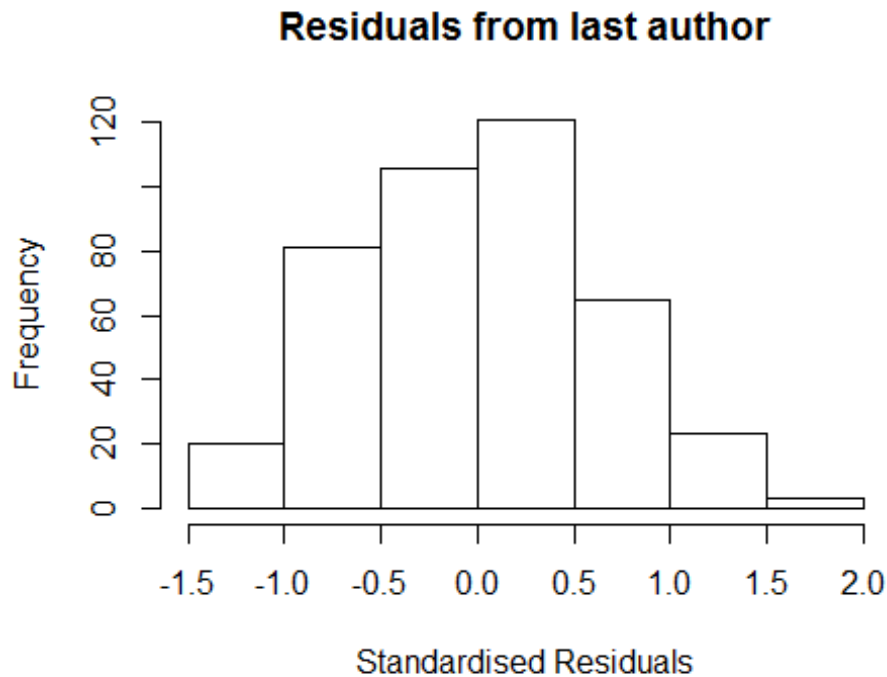
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1404"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    9    6    4    2    9    8    3    4    8   10    8    7    8    4
## 2011 2012
##   13   10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    2    3    0    3    4    3    2    7    6    6    4    5    2
## 2011 2012
##    6    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    1    3    0    3    4    3    2    6    6    5    4    5    1
## 2011 2012
##    4    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.81712059283214"
## [1] "Male first author team size 2018 geometric mean: 1.76801043048723"

## 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 = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.18224727194344"
## [1] "Male last author team size 2018 geometric mean: 1.31607401295249"

## 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 = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
##      Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
##      factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 13.837 1          3.720
## LastAuthorFemale  3.437 1          1.854
## Year              19.676 15         1.104
##
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 8 0000481870 3.27 1997 1404 5 2.729
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
```

```

## -1.15e+00 -2.14e-01 2.05e-16 2.57e-01 2.73e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.33e+00   2.54e-08  5.24e+07 < 2e-16 ***
## FirstAuthorFemale1 -3.80e-01   2.58e-01 -1.47e+00  0.14929
## LastAuthorFemale1  1.24e-01   1.14e-01  1.09e+00  0.28055
## Year1997         -7.90e-01   1.88e-01 -4.21e+00  0.00015 ***
## Year1998         -9.78e-01   1.36e-08 -7.17e+07 < 2e-16 ***
## Year1999         -2.32e-01   2.41e-01 -9.60e-01  0.34307
## Year2001          3.52e-01   1.11e-01  3.17e+00  0.00301 **
## Year2002         -7.34e-02   4.36e-01 -1.70e-01  0.86718
## Year2003         -6.03e-01   5.81e-01 -1.04e+00  0.30524
## Year2004         -3.52e-01   4.43e-01 -8.00e-01  0.43113
## Year2005         -1.83e-01   2.52e-01 -7.30e-01  0.47164
## Year2006         -8.06e-01   1.91e-01 -4.23e+00  0.00014 ***
## Year2007          2.10e-01   1.06e-01  1.98e+00  0.05460 .
## Year2008          4.22e-02   1.31e-01  3.20e-01  0.74838
## Year2009          7.19e-02   2.33e-01  3.10e-01  0.75917
## Year2010          4.30e-02   3.66e-08  1.18e+06 < 2e-16 ***
## Year2011          4.82e-03   2.16e-01  2.00e-02  0.98228
## Year2012          1.71e-01   2.69e-01  6.40e-01  0.52887
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared:  0.43, Adjusted R-squared:  0.175
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.0018);
## 6 weights are ~1. The remaining 49 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.547  0.920  0.975   0.934  0.992  0.999
## Algorithmic parameters:
##           tuning.chi              bb          tuning.psi          refine.tol
##           1.55e+00             5.00e-01          4.69e+00          1.00e-07
##           rel.tol             solve.tol          eps.outlier          eps.x
##           1.00e-07             1.00e-07          1.79e-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"

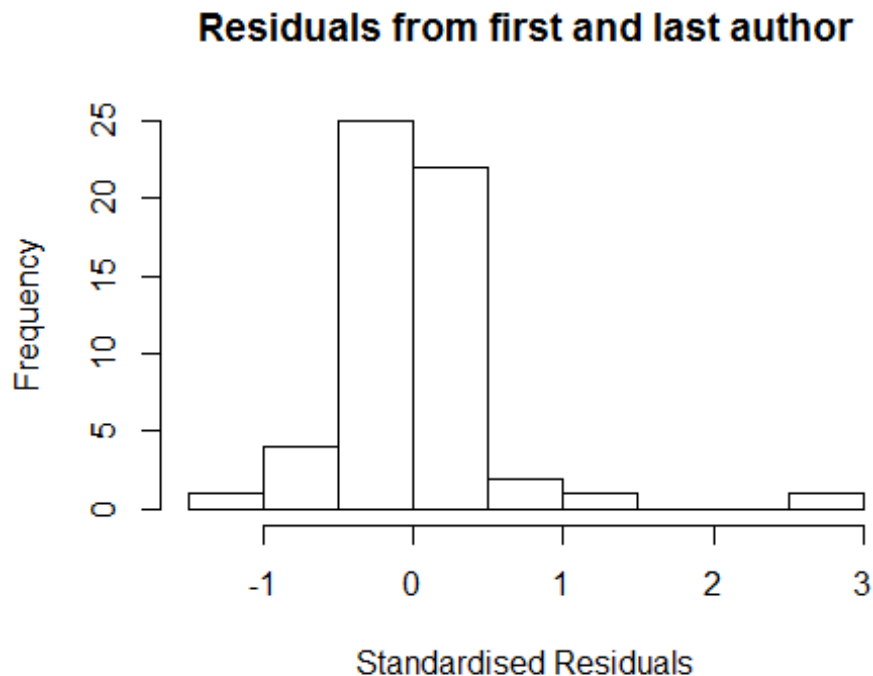
```



```
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## 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 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 8 0000481870 3.27 1997      1404      5      2.729
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16213 -0.22609 -0.00188  0.25210  2.72900
##
## Coefficients:
```

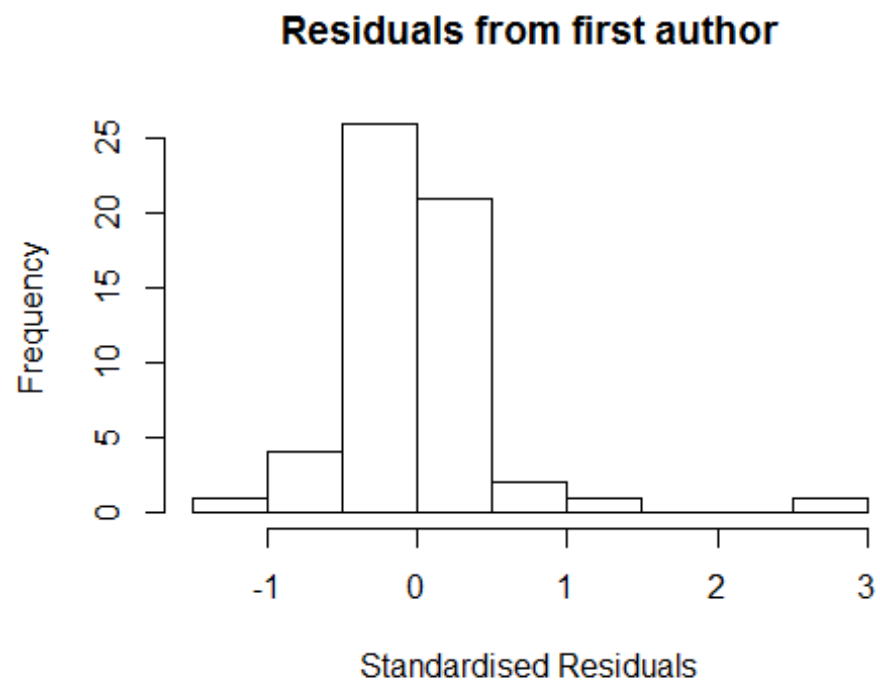
```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.33100   0.00000    Inf < 2e-16 ***
## FirstAuthorFemale1 -0.29139   0.24052   -1.21 0.23300
## Year1997          -0.79000   0.18835   -4.19 0.00015 ***
## Year1998          -0.97800   0.00000   -Inf < 2e-16 ***
## Year1999          -0.23108   0.24269   -0.95 0.34687
## Year2001           0.36363   0.11075    3.28 0.00217 **
## Year2002          -0.10085   0.44171   -0.23 0.82059
## Year2003          -0.60952   0.59704   -1.02 0.31360
## Year2004          -0.35250   0.45031   -0.78 0.43847
## Year2005          -0.16887   0.25271   -0.67 0.50791
## Year2006          -0.78139   0.18913   -4.13 0.00018 ***
## Year2007           0.24324   0.09448    2.57 0.01395 *
## Year2008           0.07491   0.13116    0.57 0.57117
## Year2009           0.02902   0.24657    0.12 0.90691
## Year2010           0.04300   0.00000    Inf < 2e-16 ***
## Year2011           0.00242   0.21108    0.01 0.99091
## Year2012           0.20643   0.27740    0.74 0.46124
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.466
## Multiple R-squared:  0.43, Adjusted R-squared:  0.197
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.0018);
## 7 weights are ~ = 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.514  0.922  0.972  0.928  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      1.79e-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"

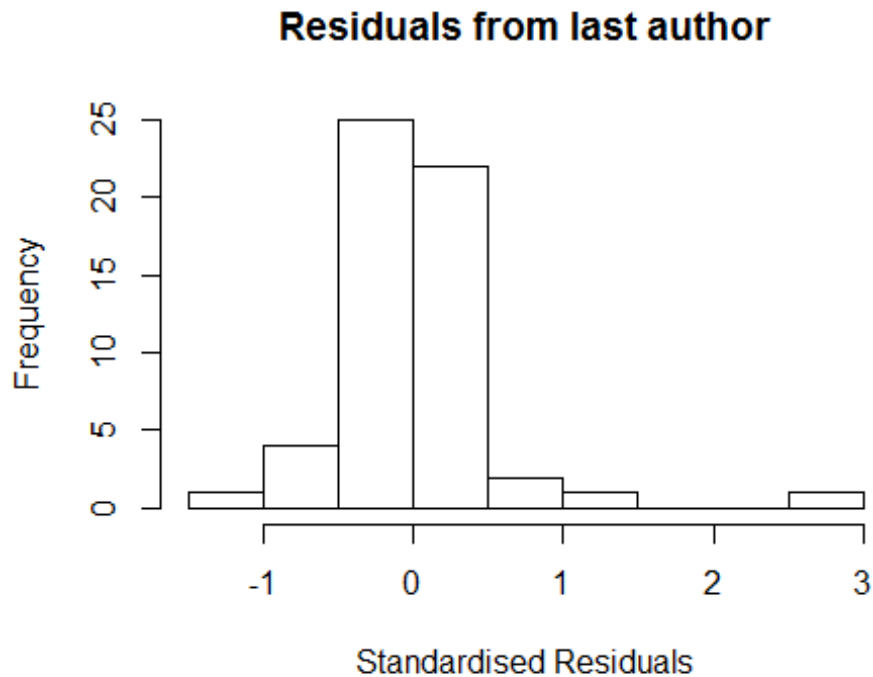
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

```

```
## 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
```



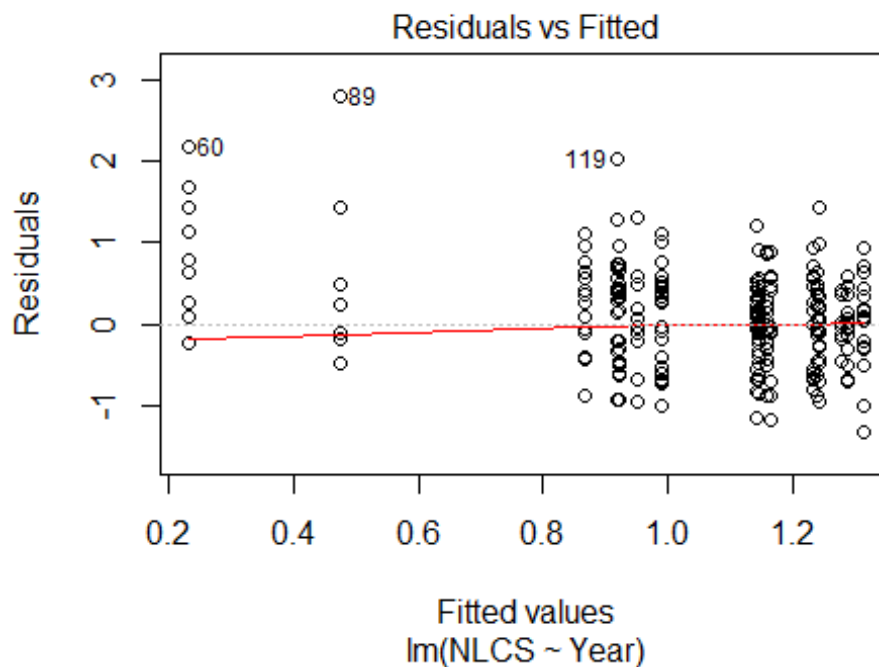
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 8 0000481870 3.27 1997      1404      5      2.729
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.09379 -0.24957  0.00163  0.25080  2.72900
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.33e+00  0.00e+00      Inf < 2e-16 ***
## LastAuthorFemale1 -5.45e-02  1.23e-01 -4.40e-01  0.65927
## Year1997       -7.90e-01  1.87e-01 -4.23e+00  0.00014 ***
## Year1998       -9.78e-01  1.58e-08 -6.17e+07 < 2e-16 ***
## Year1999       -2.32e-01  2.39e-01 -9.70e-01  0.33707
## Year2001        2.85e-01  1.19e-01  2.38e+00  0.02219 *
## Year2002       -2.76e-01  3.47e-01 -7.90e-01  0.43237
## Year2003       -5.95e-01  5.60e-01 -1.06e+00  0.29423
## Year2004       -3.52e-01  4.33e-01 -8.10e-01  0.42027
## Year2005       -2.37e-01  2.37e-01 -1.00e+00  0.32336
## Year2006       -7.83e-01  1.61e-01 -4.86e+00  1.9e-05 ***
## Year2007        2.16e-01  1.28e-01  1.68e+00  0.10026
```

```

## Year2008          8.91e-02   1.42e-01   6.30e-01   0.53314
## Year2009         -1.98e-01   1.98e-01  -1.00e+00   0.32342
## Year2010          4.30e-02   0.00e+00          Inf < 2e-16 ***
## Year2011         -1.85e-01   1.27e-01  -1.45e+00   0.15400
## Year2012         -2.94e-02   1.85e-01  -1.60e-01   0.87474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.392, Adjusted R-squared:  0.143
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.0018);
## 8 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.494  0.913  0.976  0.933  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      1.79e-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: 56"
## [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
##   67   26   22   25   13   17   14   17   26   30   18   25   22   19   18
## 2011 2012
##    8   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   52   18   21   24   12   11   13   16   20   28   15   22   17   15   16
## 2011 2012

```

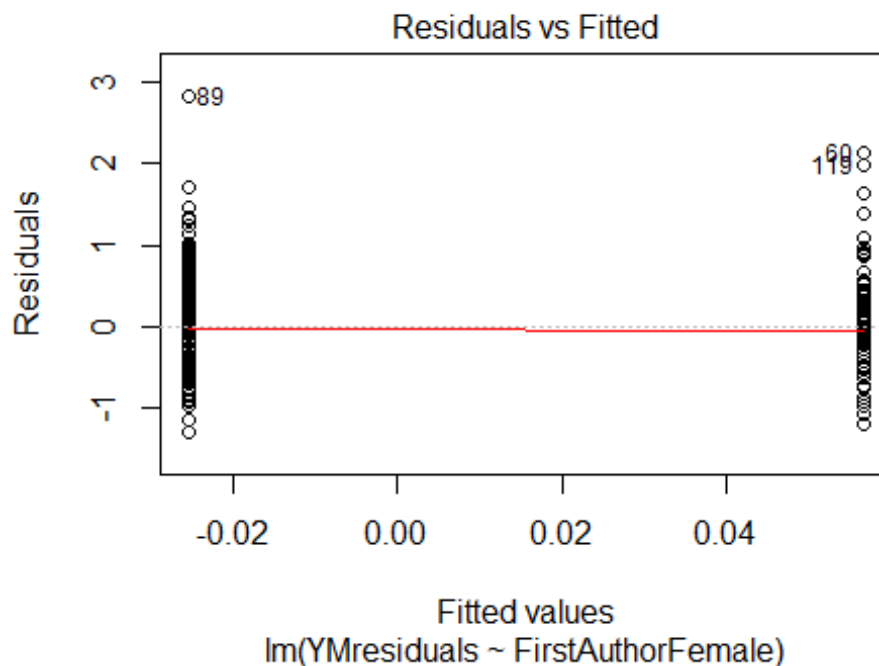
```
##      7      12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   51   16   21   24   11   11   13   15   18   25   14   20   15   15   14
## 2011 2012
##    6    12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 26, df = 16, p-value = 0.05
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0035, df = 1, p-value = 1
## [1] "Female first author team size 2018 geometric mean: 2.21336383940064"
## [1] "Male first author team size 2018 geometric mean: 1.55561667808903"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

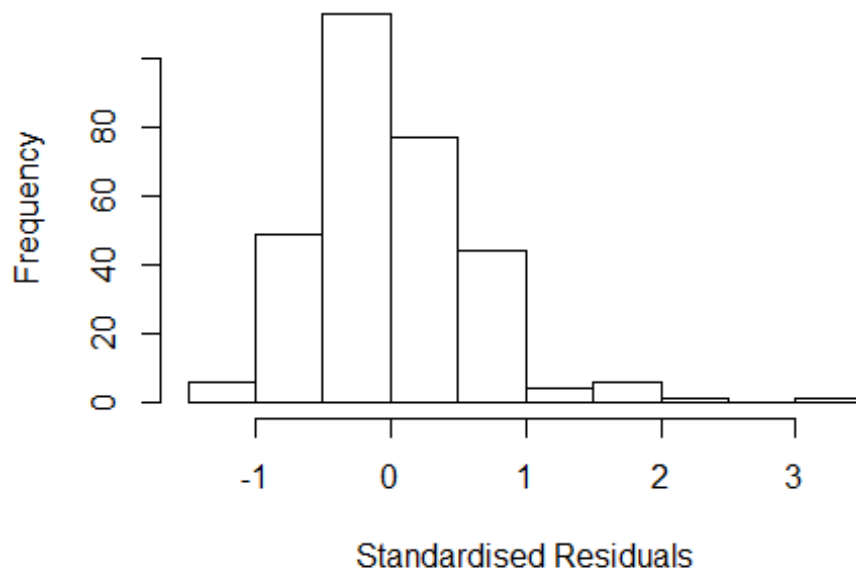
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 38, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.35215804504935"
## [1] "Male last author team size 2018 geometric mean: 1.47898133590682"

## 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 = 48, 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 4.725 1      2.174
## LastAuthorFemale  4.051 1      2.013
## UniqueAuthors    7.544 3      1.400
## Year             11.880 16     1.080
```

Residuals from first and last author and team size



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 89 0000481870 3.27 1997      1404      5      3.212
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2038 -0.2895 -0.0377  0.4129  3.2116
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0377    0.0364   1.04  0.30064
## FirstAuthorFemale1 -0.0934    0.1302  -0.72  0.47392
## LastAuthorFemale1  0.1557    0.1370   1.14  0.25678
## UniqueAuthors2    0.1830    0.1126   1.62  0.10530
## UniqueAuthors3    0.2423    0.1397   1.73  0.08400 .
## UniqueAuthors4    0.0831    0.2413   0.34  0.73067
## Year1997          0.0207    0.0676   0.31  0.76024
## Year1998          0.6404    0.2398   2.67  0.00801 **
## Year1999          0.9224    0.1902   4.85  2.1e-06 ***
## Year2000          0.6892    0.1926   3.58  0.00041 ***
```

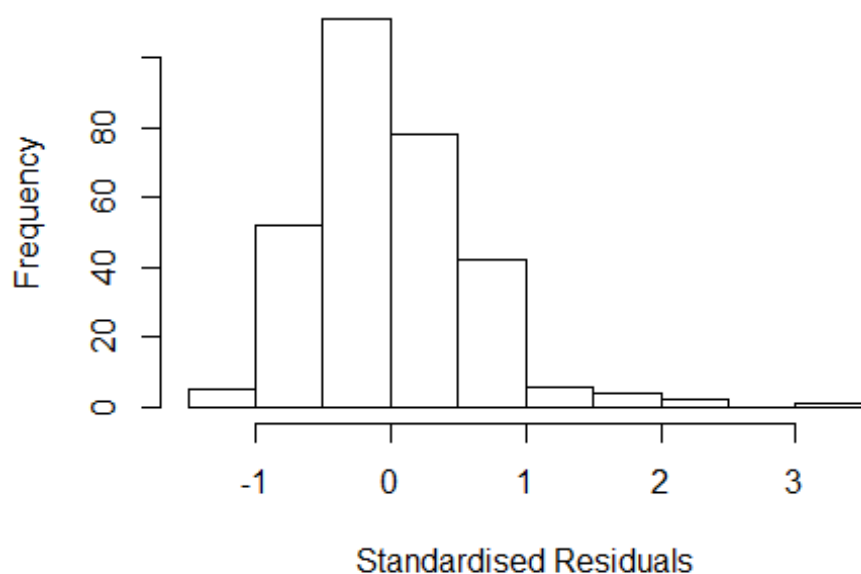


```

## Year2001          1.1163      0.2598      4.30  2.4e-05 ***
## Year2002          0.9981      0.1442      6.92  3.1e-11 ***
## Year2003          0.9848      0.2771      3.55  0.00045 ***
## Year2004          0.8065      0.2263      3.56  0.00043 ***
## Year2005          1.0151      0.1187      8.55  8.1e-16 ***
## Year2006          0.9527      0.1687      5.65  4.0e-08 ***
## Year2007          0.7478      0.2075      3.60  0.00037 ***
## Year2008          1.2283      0.1213     10.13 < 2e-16 ***
## Year2009          1.0404      0.1861      5.59  5.4e-08 ***
## Year2010          1.1661      0.1935      6.03  5.3e-09 ***
## Year2011          1.0712      0.1565      6.85  4.8e-11 ***
## Year2012          1.0729      0.1487      7.21  5.2e-12 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.432
## Multiple R-squared:  0.472, Adjusted R-squared:  0.433
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 2 observations c(47,64) are outliers with |weight| = 0 ( < 0.00033);
## 51 weights are ~= 1. The remaining 248 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0033 0.7840 0.8990 0.8460 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          3.32e-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 3.414 1          1.848
## LastAuthorFemale 3.133 1          1.770
## Year              2.689 16          1.031

```

Residuals from first and last author

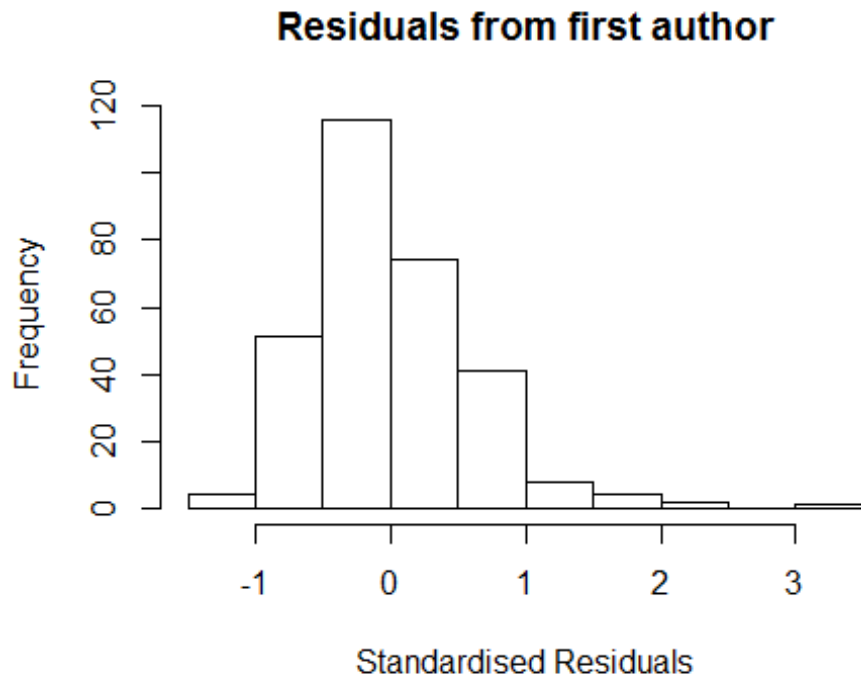


```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 89 0000481870 3.27 1997      1404      5      3.181
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3450 -0.2498 -0.0375  0.3841  3.1810
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.0375     0.0362   1.04  0.30147
## FirstAuthorFemale1 -0.0380     0.1085  -0.35  0.72673
## LastAuthorFemale1  0.1181     0.1229   0.96  0.33702
## Year1997          0.0515     0.0657   0.78  0.43331
## Year1998          0.7176     0.2457   2.92  0.00377 **
## Year1999          0.9474     0.1852   5.12  5.8e-07 ***
## Year2000          0.8004     0.1683   4.76  3.1e-06 ***
## Year2001          1.1998     0.2622   4.58  7.1e-06 ***
## Year2002          1.0967     0.1437   7.63  3.6e-13 ***
## Year2003          1.1275     0.2816   4.00  8.0e-05 ***
## Year2004          0.8722     0.2450   3.56  0.00043 ***
## Year2005          1.0979     0.1088  10.10 < 2e-16 ***
```

```

## Year2006          1.0976      0.1493      7.35  2.1e-12 ***
## Year2007          0.8233      0.1780      4.63  5.7e-06 ***
## Year2008          1.2450      0.1176     10.59 < 2e-16 ***
## Year2009          1.1422      0.1765      6.47  4.3e-10 ***
## Year2010          1.3075      0.1478      8.84 < 2e-16 ***
## Year2011          1.1787      0.1244      9.48 < 2e-16 ***
## Year2012          1.1812      0.1526      7.74  1.8e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.434
## Multiple R-squared:  0.462, Adjusted R-squared:  0.428
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 3 observations c(47,64,85) are outliers with |weight| = 0 ( < 0.00033);
## 50 weights are ~= 1. The remaining 248 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.024  0.776   0.908   0.847   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      3.32e-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 1.552  1         1.246
## Year              1.552 16         1.014

```

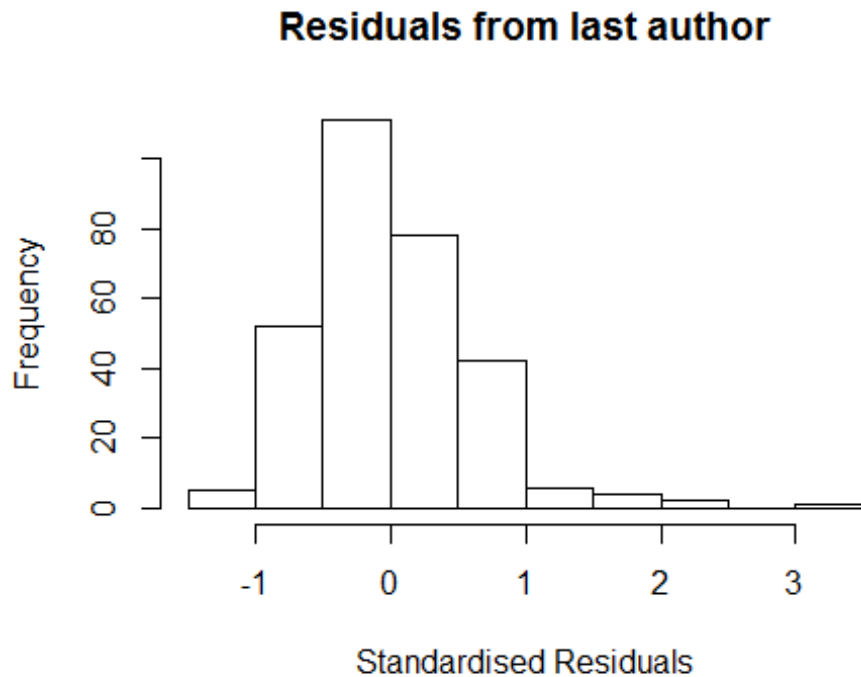


```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 89 0000481870 3.27 1997      1404      5      3.181
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3758 -0.2268 -0.0396  0.3824  3.1884
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.0396     0.0332   1.19  0.23479
## FirstAuthorFemale1 0.0347     0.0734   0.47  0.63685
## Year1997          0.0420     0.0595   0.71  0.48103
## Year1998          0.7493     0.2697   2.78  0.00583 **
## Year1999          0.9510     0.1887   5.04  8.3e-07 ***
## Year2000          0.7948     0.1619   4.91  1.5e-06 ***
## Year2001          1.2154     0.2771   4.39  1.6e-05 ***
## Year2002          1.0850     0.1451   7.48  9.5e-13 ***
## Year2003          1.1576     0.3005   3.85  0.00015 ***
## Year2004          0.9029     0.2465   3.66  0.00030 ***
## Year2005          1.1133     0.1050  10.60 < 2e-16 ***
## Year2006          1.1110     0.1490   7.46  1.1e-12 ***
```

```

## Year2007          0.8187      0.1749      4.68  4.5e-06 ***
## Year2008          1.2567      0.1160     10.84 < 2e-16 ***
## Year2009          1.1069      0.1637      6.76  7.7e-11 ***
## Year2010          1.3363      0.1446      9.24 < 2e-16 ***
## Year2011          1.1471      0.1233      9.31 < 2e-16 ***
## Year2012          1.2042      0.1524      7.90  6.1e-14 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.417
## Multiple R-squared:  0.472, Adjusted R-squared:  0.441
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 3 observations c(47,64,85) are outliers with |weight| = 0 ( < 0.00033);
## 50 weights are ~= 1. The remaining 248 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0078 0.7570 0.8990 0.8360 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      3.32e-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 1.544 1      1.243
## Year      1.544 16      1.014

```



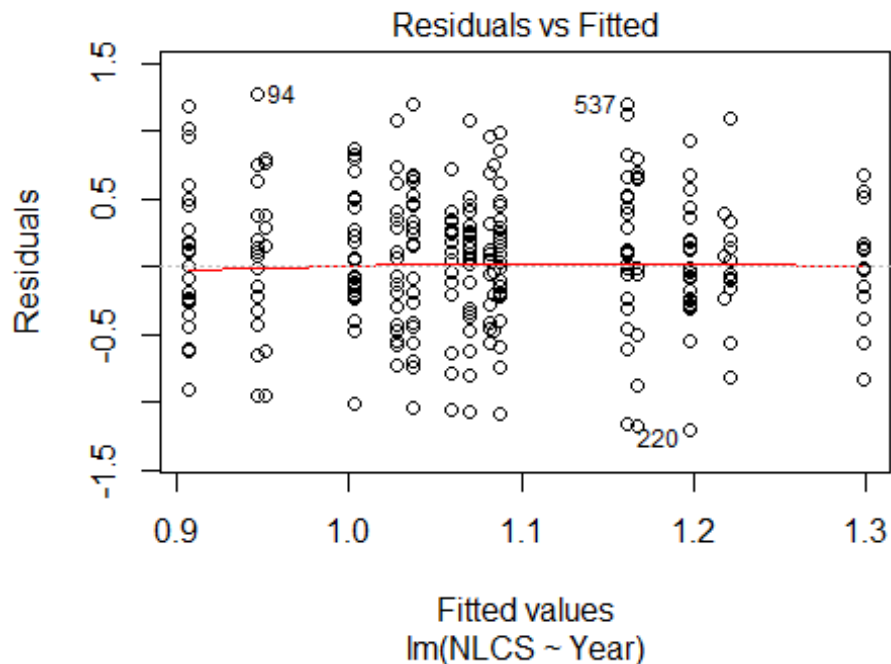
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 89 0000481870 3.27 1997      1404      5      3.181
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3482 -0.2484 -0.0342  0.3808  3.1894
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.0342     0.0346   0.99  0.32390
## LastAuthorFemale1 0.0912     0.0874   1.04  0.29754
## Year1997        0.0464     0.0624   0.74  0.45740
## Year1998        0.7218     0.2537   2.85  0.00476 **
## Year1999        0.9468     0.1852   5.11  5.8e-07 ***
## Year2000        0.7950     0.1649   4.82  2.3e-06 ***
## Year2001        1.1982     0.2641   4.54  8.4e-06 ***
## Year2002        1.0920     0.1427   7.65  3.1e-13 ***
## Year2003        1.1326     0.2847   3.98  8.8e-05 ***
## Year2004        0.8768     0.2451   3.58  0.00041 ***
## Year2005        1.0967     0.1084  10.11 < 2e-16 ***
## Year2006        1.1001     0.1486   7.40  1.5e-12 ***
```

```

## Year2007          0.8158      0.1795      4.54  8.2e-06 ***
## Year2008          1.2458      0.1195     10.43 < 2e-16 ***
## Year2009          1.1257      0.1608      7.00  1.8e-11 ***
## Year2010          1.3140      0.1485      8.85 < 2e-16 ***
## Year2011          1.1603      0.1099     10.56 < 2e-16 ***
## Year2012          1.1768      0.1524      7.72  2.0e-13 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.429
## Multiple R-squared:  0.465, Adjusted R-squared:  0.433
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 3 observations c(47,64,85) are outliers with |weight| = 0 ( < 0.00033);
## 50 weights are ~= 1. The remaining 248 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0178 0.7700 0.9060 0.8440 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      3.32e-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: 301"
## [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
##   10   19   8    5   16   22   22   16   12   29   21   30   30   32   41
## 2011 2012
##   35   29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9   11   5    4   13   18   17   14   12   25   20   26   23   24   31

```

```
## 2011 2012
## 28 24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 11 5 4 13 15 16 13 12 22 16 22 21 23 28
## 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 = 16, df = 16, p-value = 0.4
```

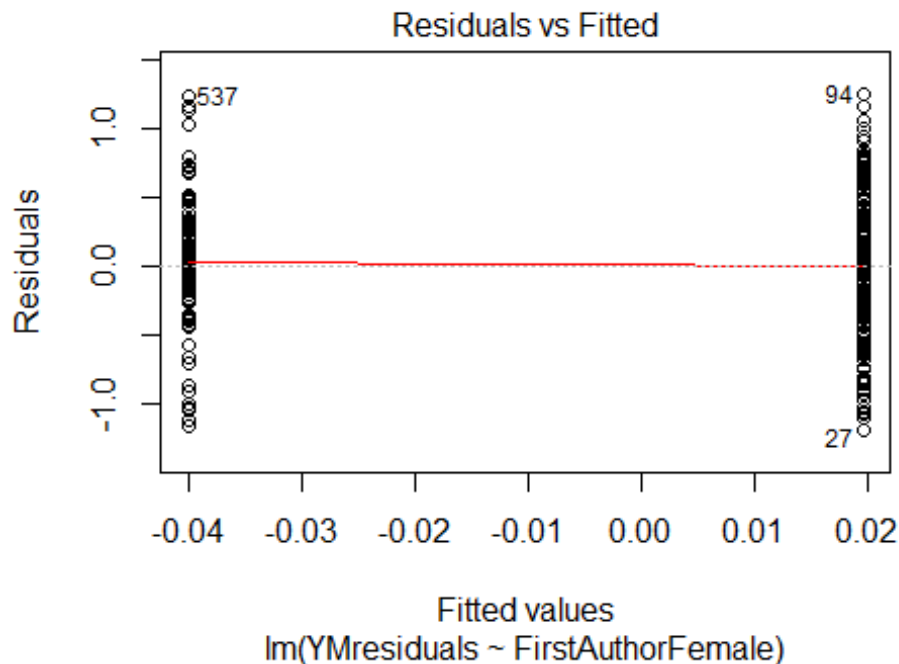


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.042, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 2.17665617921286"
## [1] "Male first author team size 2018 geometric mean: 1.73552611877147"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



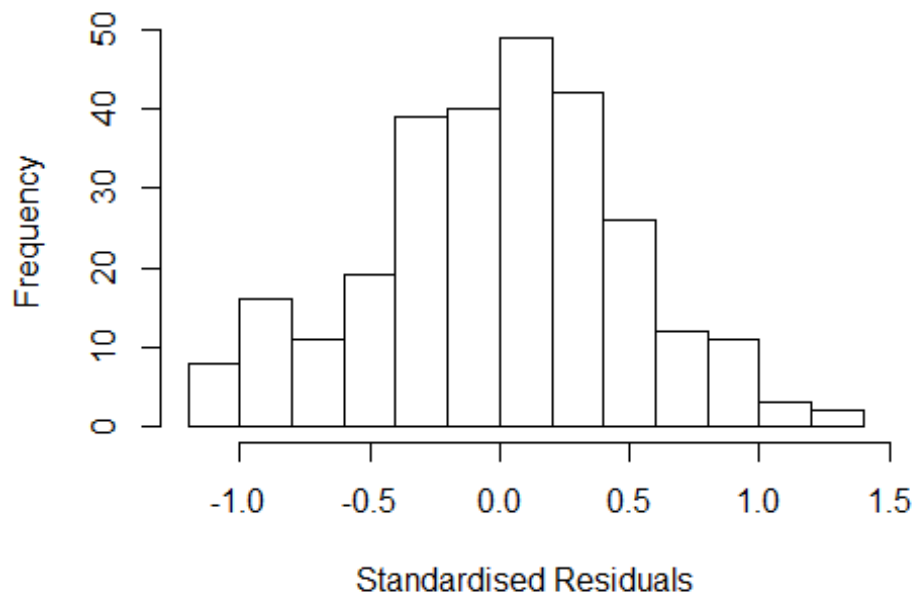
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.00762084174617"
## [1] "Male last author team size 2018 geometric mean: 1.91233913378706"

## 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 = 95, 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 1.919  1      1.385
## LastAuthorFemale  1.390  1      1.179
## UniqueAuthors    2.627  3      1.175
## Year              4.378 16      1.047
```

Residuals from first and last author and team size



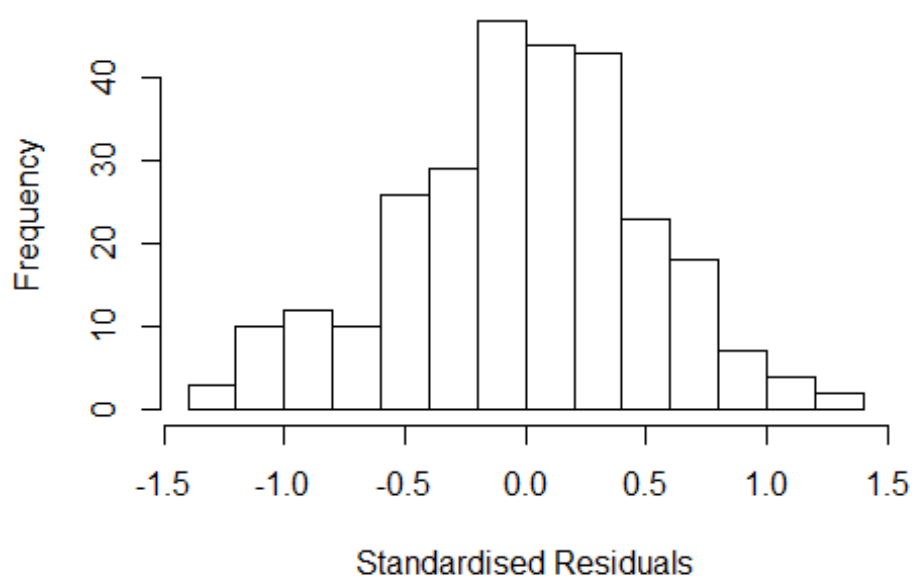
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1548 -0.3194 0.0192 0.3027 1.2750
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9755 0.2635 3.70 0.00026 ***
## FirstAuthorFemale1 -0.1570 0.0763 -2.06 0.04060 *
## LastAuthorFemale1 0.1424 0.0685 2.08 0.03846 *
## UniqueAuthors2 0.1557 0.0764 2.04 0.04243 *
## UniqueAuthors3 -0.0206 0.1145 -0.18 0.85755
## UniqueAuthors4 0.4655 0.1320 3.53 0.00050 ***
## Year1997 0.1793 0.3989 0.45 0.65354
## Year1998 -0.0155 0.2986 -0.05 0.95855
## Year1999 0.0834 0.3187 0.26 0.79386
## Year2000 0.2496 0.2867 0.87 0.38492
```

```

## Year2001          -0.1862      0.2849   -0.65   0.51407
## Year2002           0.0998      0.2875    0.35   0.72881
## Year2003           0.0489      0.3041    0.16   0.87250
## Year2004           0.1671      0.2870    0.58   0.56105
## Year2005           0.1850      0.2744    0.67   0.50082
## Year2006          -0.0594      0.2978   -0.20   0.84196
## Year2007          -0.1585      0.2869   -0.55   0.58113
## Year2008           0.1097      0.2767    0.40   0.69216
## Year2009           0.0314      0.3038    0.10   0.91778
## Year2010          -0.0525      0.2820   -0.19   0.85242
## Year2011           0.1043      0.2809    0.37   0.71078
## Year2012           0.0431      0.3056    0.14   0.88785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.467
## Multiple R-squared:  0.105, Adjusted R-squared:  0.0314
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 256 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.437  0.860  0.951  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      3.60e-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 1.806 1      1.344
## LastAuthorFemale  1.315 1      1.147
## Year              1.891 16      1.020

```

Residuals from first and last author



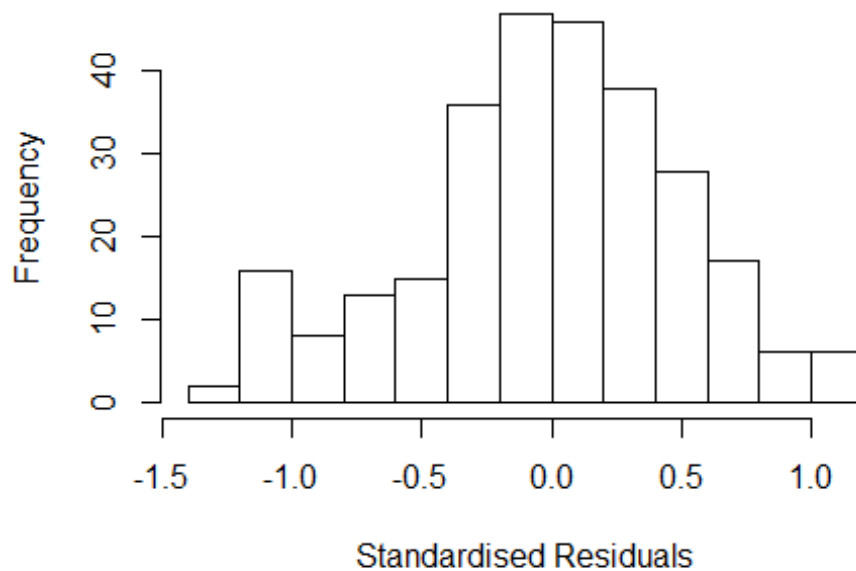
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.29294 -0.29059 0.00946 0.31366 1.25075
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01545 0.27032 3.76 0.00021 ***
## FirstAuthorFemale1 -0.14918 0.07753 -1.92 0.05543 .
## LastAuthorFemale1 0.16928 0.06645 2.55 0.01143 *
## Year1997 0.27749 0.43571 0.64 0.52478
## Year1998 -0.02006 0.31537 -0.06 0.94933
## Year1999 0.10522 0.31176 0.34 0.73600
## Year2000 0.30255 0.29475 1.03 0.30564
## Year2001 -0.21549 0.29404 -0.73 0.46431
## Year2002 0.11590 0.29574 0.39 0.69546
## Year2003 0.03841 0.30924 0.12 0.90125
## Year2004 0.19399 0.29458 0.66 0.51078
## Year2005 0.19192 0.28492 0.67 0.50117
```

```

## Year2006          0.00741    0.31377    0.02  0.98117
## Year2007         -0.13270    0.29255   -0.45  0.65050
## Year2008          0.09709    0.28521    0.34  0.73382
## Year2009          0.02740    0.31451    0.09  0.93064
## Year2010         -0.02092    0.29112   -0.07  0.94275
## Year2011          0.10144    0.28962    0.35  0.72644
## Year2012          0.06590    0.31229    0.21  0.83303
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared:  0.0753, Adjusted R-squared:  0.011
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 255 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.448  0.854  0.954  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      3.60e-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 1.695 1      1.302
## Year              1.695 16      1.017

```

Residuals from first author



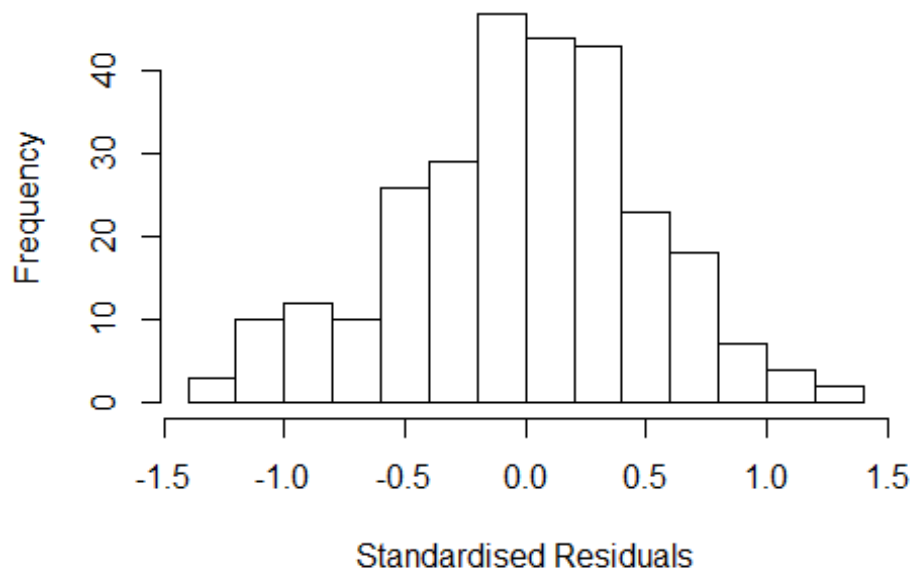
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2912 -0.3047  0.0201  0.3049  1.1770
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.02027    0.27871     3.66  0.0003 ***
## FirstAuthorFemale1 -0.07274    0.07641    -0.95  0.3420
## Year1997          0.27088    0.44601     0.61  0.5441
## Year1998          0.04908    0.33140     0.15  0.8824
## Year1999          0.20872    0.31778     0.66  0.5119
## Year2000          0.29776    0.30484     0.98  0.3296
## Year2001         -0.18600    0.30191    -0.62  0.5384
## Year2002          0.12353    0.30163     0.41  0.6825
## Year2003          0.03993    0.31416     0.13  0.8990
## Year2004          0.19258    0.30265     0.64  0.5251
## Year2005          0.25860    0.29394     0.88  0.3798
## Year2006          0.06520    0.32312     0.20  0.8402
```

```

## Year2007          -0.10529    0.30137   -0.35    0.7271
## Year2008          0.11300    0.29439    0.38    0.7014
## Year2009          0.04726    0.32356    0.15    0.8840
## Year2010         -0.00975    0.29833   -0.03    0.9739
## Year2011          0.13282    0.29726    0.45    0.6554
## Year2012          0.11182    0.31882    0.35    0.7261
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.0593, Adjusted R-squared:  -0.00218
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 253 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.428  0.840   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      3.60e-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

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2939 -0.3215 0.0225 0.3068 1.2245
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.020595 0.278928 3.66 0.00031 ***
## LastAuthorFemale1 0.103119 0.064014 1.61 0.10842
## Year1997 0.224983 0.425915 0.53 0.59779
## Year1998 -0.018390 0.334385 -0.05 0.95618
## Year1999 0.116322 0.309579 0.38 0.70741
## Year2000 0.277387 0.303401 0.91 0.36143
## Year2001 -0.240318 0.303273 -0.79 0.42884
## Year2002 0.072380 0.298245 0.24 0.80844
## Year2003 0.000923 0.315593 0.00 0.99767
## Year2004 0.168036 0.302750 0.56 0.57935
## Year2005 0.170209 0.293368 0.58 0.56229
## Year2006 -0.030788 0.321771 -0.10 0.92385
```

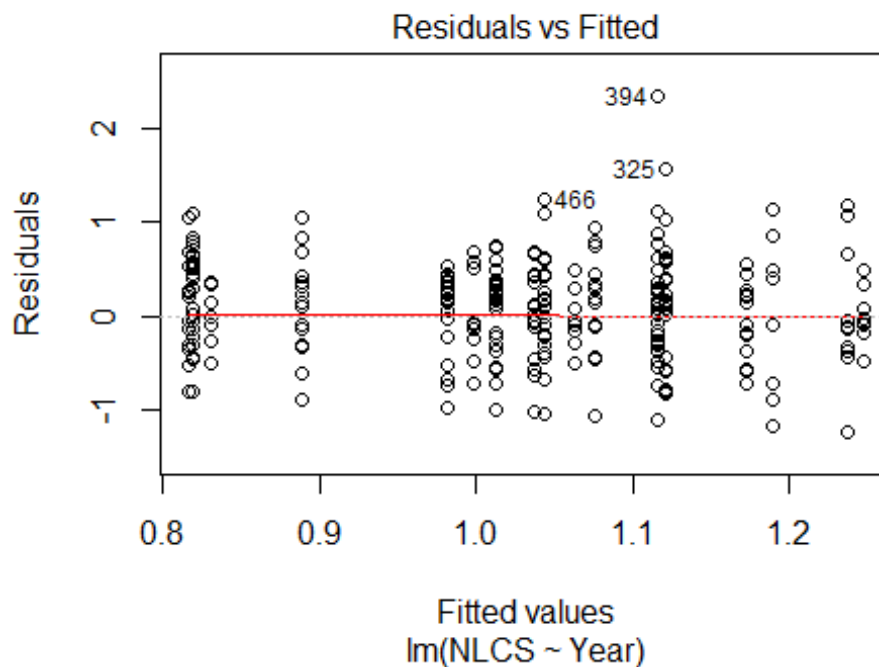


```

## Year2007          -0.138869    0.303896    -0.46    0.64808
## Year2008           0.063440    0.292521     0.22    0.82848
## Year2009          -0.016123    0.321105    -0.05    0.95999
## Year2010          -0.078765    0.295179    -0.27    0.78981
## Year2011           0.036866    0.294113     0.13    0.90035
## Year2012           0.071986    0.323313     0.22    0.82398
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.468
## Multiple R-squared:  0.0649, Adjusted R-squared:  0.00373
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 256 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.424  0.847   0.952   0.891   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      3.60e-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: 278"
## [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
##   13   11    9   20    9   15   12   22   11   23   19   31   24   34   38
## 2011 2012
##   42   30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12    9    9   20    9   13    8   17    9   18   17   27   17   28   29
## 2011 2012

```

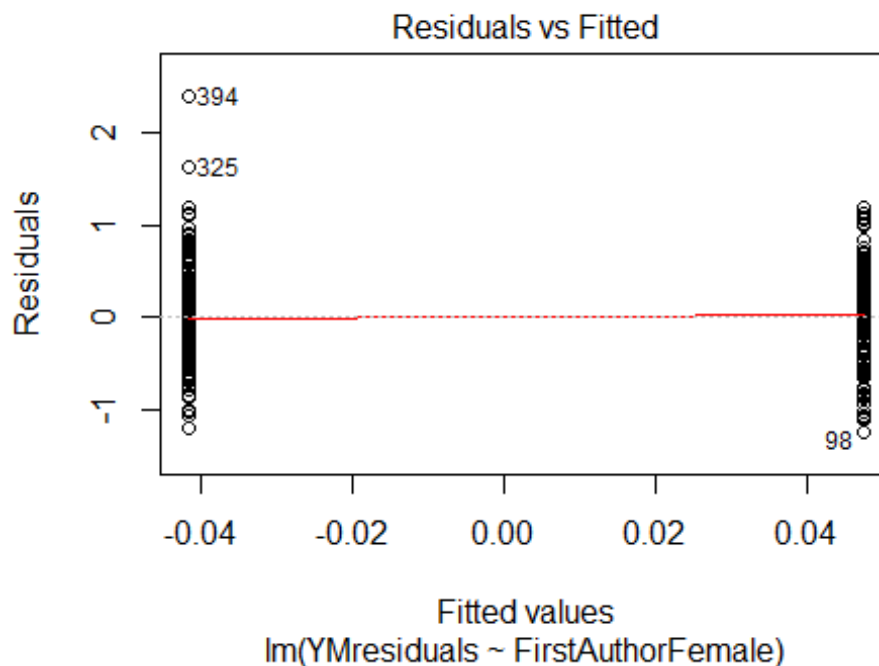
```
## 35 24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 7 8 19 9 11 7 15 8 17 16 26 16 27 28
## 2011 2012
## 33 23
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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: 2.28942848510666"
## [1] "Male first author team size 2018 geometric mean: 1.46281454322389"
## 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 = 53, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.13982638786733"
## [1] "Male last author team size 2018 geometric mean: 1.5130857494229"

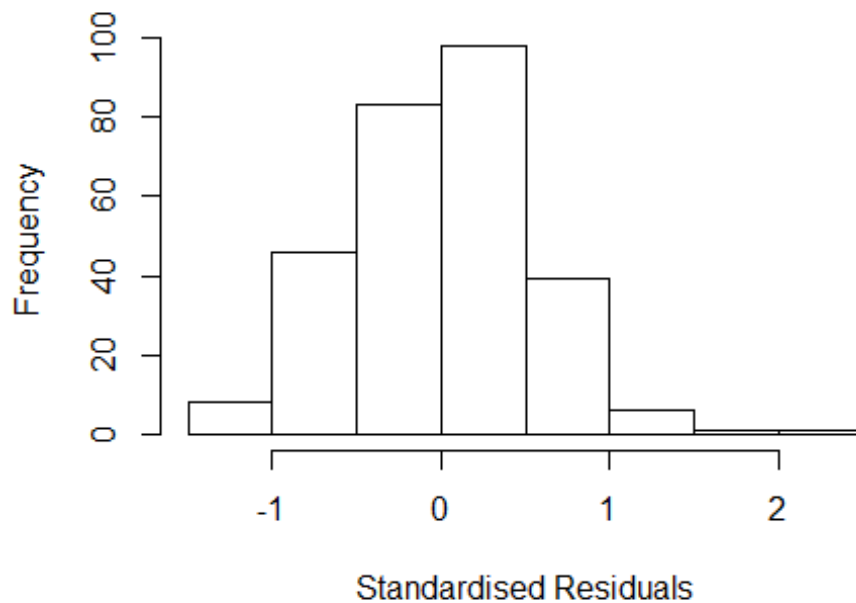
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 49, 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	5.113	1	2.261
LastAuthorFemale	5.525	1	2.351
UniqueAuthors	51.395	4	1.636
Year	77.211	16	1.145

Residuals from first and last author and team size



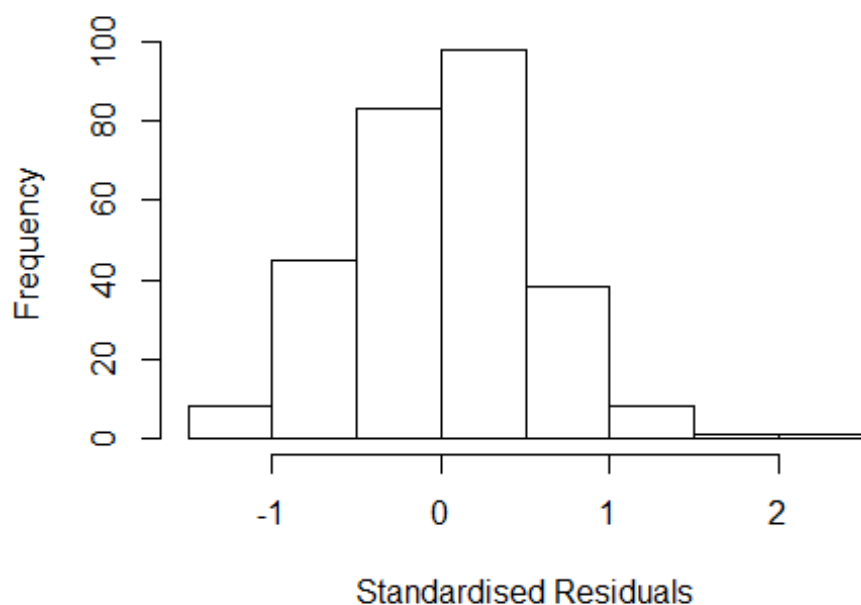
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.293 -0.366 0.017 0.372 2.086
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06696 0.21288 5.01 1e-06 ***
## FirstAuthorFemale1 0.00707 0.09790 0.07 0.942
## LastAuthorFemale1 0.16450 0.10617 1.55 0.123
## UniqueAuthors2 0.09603 0.07985 1.20 0.230
## UniqueAuthors3 0.24060 0.12834 1.87 0.062 .
## UniqueAuthors4 0.23491 0.19532 1.20 0.230
## UniqueAuthors5 0.15655 0.11993 1.31 0.193
## Year1997 -0.32518 0.24833 -1.31 0.192
## Year1998 0.12257 0.23639 0.52 0.605
## Year1999 -0.20076 0.23795 -0.84 0.400
```

```

## Year2000      -0.26680    0.24973   -1.07    0.286
## Year2001      -0.32397    0.29680   -1.09    0.276
## Year2002       0.05454    0.91059    0.06    0.952
## Year2003      -0.22215    0.25035   -0.89    0.376
## Year2004      -0.17486    0.22334   -0.78    0.434
## Year2005      -0.18731    0.23890   -0.78    0.434
## Year2006      -0.06082    0.22138   -0.27    0.784
## Year2007      -0.25208    0.22263   -1.13    0.259
## Year2008      -0.13334    0.28696   -0.46    0.643
## Year2009      -0.18446    0.23191   -0.80    0.427
## Year2010      -0.43775    0.23668   -1.85    0.066 .
## Year2011      -0.08636    0.22386   -0.39    0.700
## Year2012      -0.26693    0.22910   -1.17    0.245
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.0943, Adjusted R-squared:  0.0174
## Convergence in 33 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 261 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0869 0.8780 0.9500 0.9060 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          3.55e-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 1.882 1          1.372
## LastAuthorFemale  2.162 1          1.470
## Year              1.775 16          1.018

```

Residuals from first and last author

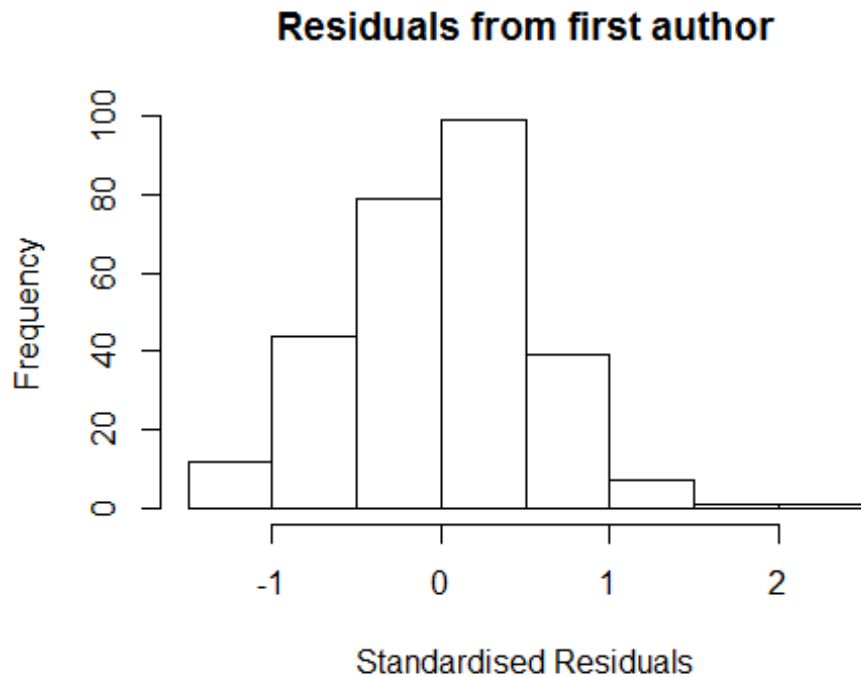


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4321 -0.3510 0.0119 0.3718 2.2464
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.11205 0.19930 5.58 6e-08 ***
## FirstAuthorFemale1 0.00888 0.08798 0.10 0.920
## LastAuthorFemale1 0.17627 0.09201 1.92 0.056 .
## Year1997 -0.37407 0.23716 -1.58 0.116
## Year1998 0.09750 0.22587 0.43 0.666
## Year1999 -0.22104 0.22454 -0.98 0.326
## Year2000 -0.18843 0.25185 -0.75 0.455
## Year2001 -0.35056 0.28769 -1.22 0.224
## Year2002 0.13492 0.72038 0.19 0.852
## Year2003 -0.25227 0.24309 -1.04 0.300
## Year2004 -0.20334 0.21326 -0.95 0.341
## Year2005 -0.13717 0.23473 -0.58 0.559
```

```

## Year2006      -0.05737    0.21045   -0.27    0.785
## Year2007      -0.21742    0.21345   -1.02    0.309
## Year2008      -0.13030    0.27513   -0.47    0.636
## Year2009      -0.15167    0.22826   -0.66    0.507
## Year2010      -0.41268    0.22900   -1.80    0.073 .
## Year2011      -0.06875    0.21629   -0.32    0.751
## Year2012      -0.21961    0.22173   -0.99    0.323
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.556
## Multiple R-squared:  0.0745, Adjusted R-squared:  0.0112
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 258 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0658 0.8730 0.9530 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      3.55e-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 1.156 1      1.075
## Year      1.156 16      1.005

```



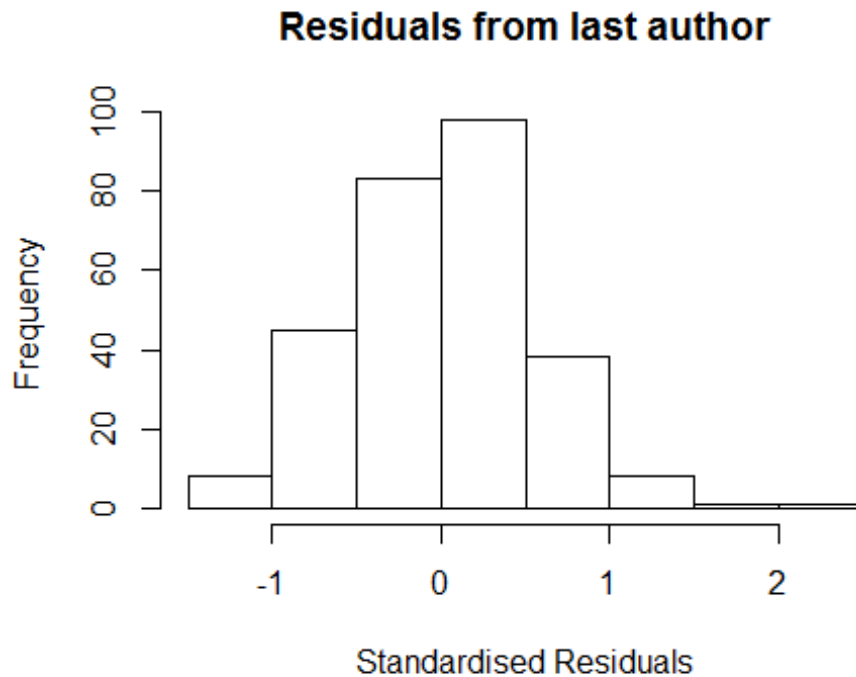
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2978 -0.3633 0.0246 0.3667 2.4192
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1398 0.1970 5.78 2.1e-08 ***
## FirstAuthorFemale1 0.1242 0.0696 1.79 0.075 .
## Year1997 -0.3856 0.2311 -1.67 0.096 .
## Year1998 0.0734 0.2264 0.32 0.746
## Year1999 -0.2004 0.2270 -0.88 0.378
## Year2000 -0.2057 0.2480 -0.83 0.408
## Year2001 -0.3499 0.2840 -1.23 0.219
## Year2002 0.0337 0.6233 0.05 0.957
## Year2003 -0.2577 0.2447 -1.05 0.293
## Year2004 -0.2002 0.2118 -0.94 0.346
## Year2005 -0.1299 0.2326 -0.56 0.577
## Year2006 -0.0532 0.2105 -0.25 0.801
```



```

## Year2007          -0.2178      0.2133    -1.02     0.308
## Year2008          -0.0837      0.2663    -0.31     0.753
## Year2009          -0.1629      0.2293    -0.71     0.478
## Year2010          -0.3944      0.2351    -1.68     0.095 .
## Year2011          -0.0930      0.2140    -0.43     0.664
## Year2012          -0.2203      0.2195    -1.00     0.317
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0593, Adjusted R-squared:  -0.00132
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 256 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.016  0.867   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      3.55e-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 1.289 1          1.135
## Year              1.289 16          1.008

```



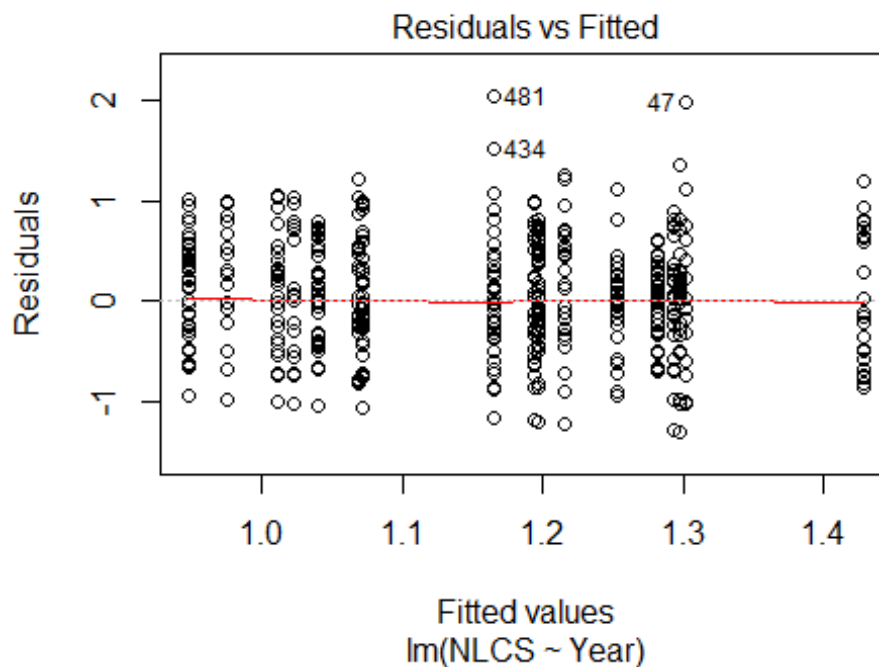
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4375 -0.3503 0.0112 0.3717 2.2386
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1128 0.1982 5.61 5e-08 ***
## LastAuthorFemale1 0.1822 0.0716 2.54 0.012 *
## Year1997 -0.3741 0.2366 -1.58 0.115
## Year1998 0.0984 0.2256 0.44 0.663
## Year1999 -0.2212 0.2229 -0.99 0.322
## Year2000 -0.1874 0.2524 -0.74 0.459
## Year2001 -0.3500 0.2878 -1.22 0.225
## Year2002 0.1424 0.7150 0.20 0.842
## Year2003 -0.2520 0.2432 -1.04 0.301
## Year2004 -0.2027 0.2132 -0.95 0.343
## Year2005 -0.1372 0.2344 -0.59 0.559
## Year2006 -0.0566 0.2104 -0.27 0.788
```

```

## Year2007          -0.2156      0.2136   -1.01    0.314
## Year2008          -0.1309      0.2738   -0.48    0.633
## Year2009          -0.1500      0.2279   -0.66    0.511
## Year2010          -0.4133      0.2283   -1.81    0.071 .
## Year2011          -0.0676      0.2162   -0.31    0.755
## Year2012          -0.2186      0.2221   -0.98    0.326
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.0748, Adjusted R-squared:  0.0152
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 258 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0648 0.8720 0.9530 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      3.55e-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: 282"
## [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
##   36   28   35   31   27   26   34   35   28   40   41   55   54   42   54
## 2011 2012
##   55   45
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   26   17   26   27   23   19   24   28   22   30   39   48   38   32   44
## 2011 2012

```

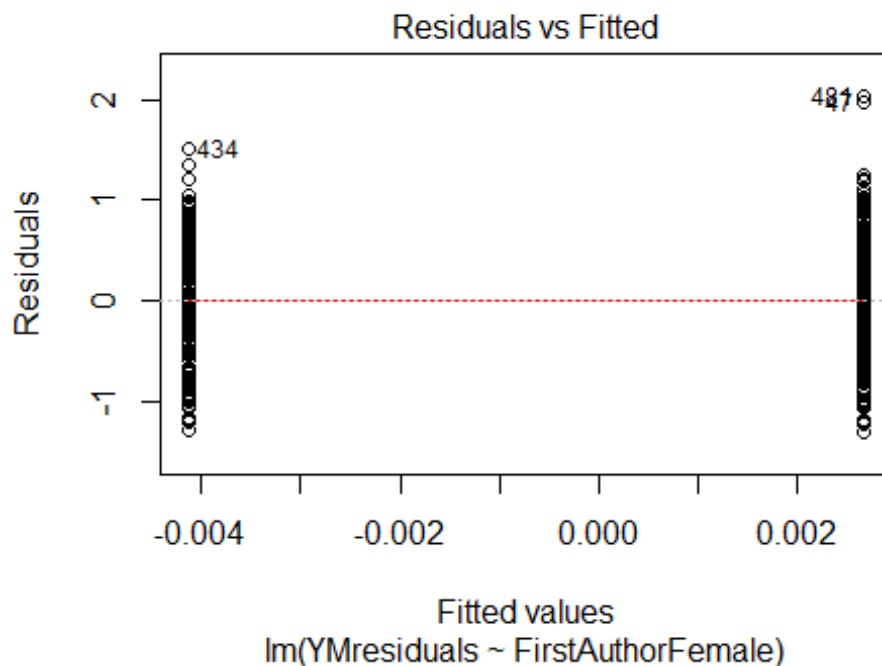
```
## 45 38
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 25 15 24 27 22 18 23 27 22 29 35 42 35 31 42
## 2011 2012
## 42 35
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 0.61, df = 1, p-value = 0.4
## [1] "Female first author team size 2018 geometric mean: 1.99452162202948"
## [1] "Male first author team size 2018 geometric mean: 2.08994304130631"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

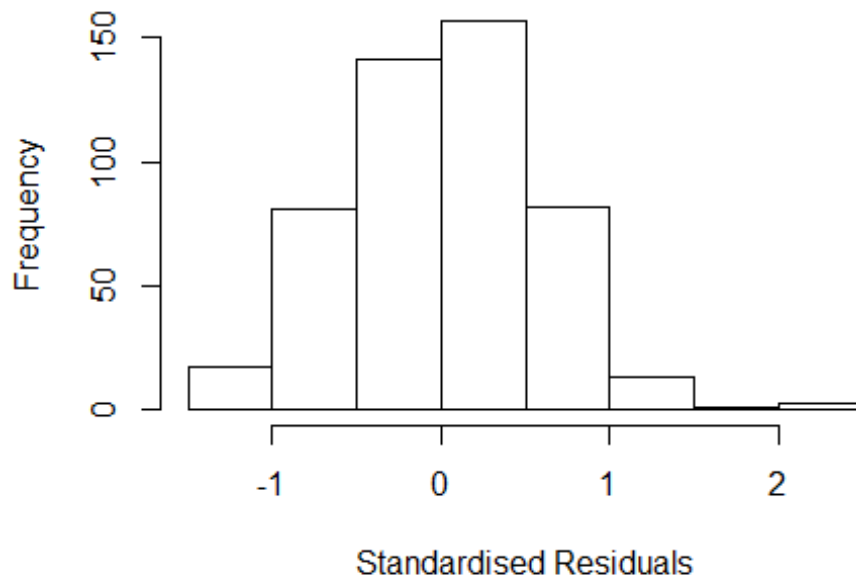
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 220, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.08012471310386"
## [1] "Male last author team size 2018 geometric mean: 2.01674700568984"

## 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 = 0.9
## 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 1.860  1      1.364
## LastAuthorFemale  1.646  1      1.283
## UniqueAuthors    2.241  4      1.106
## Year              2.280 16      1.026
```

Residuals from first and last author and team size



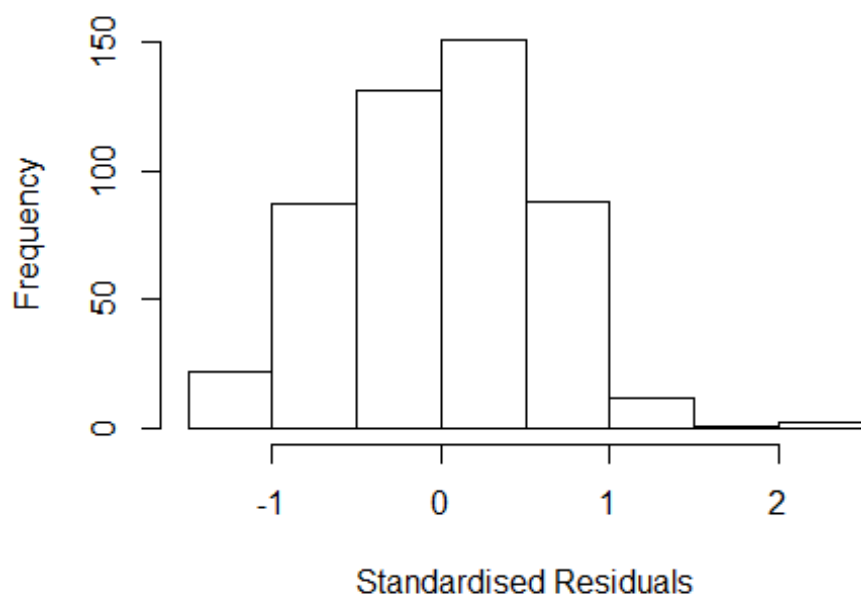
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4473 -0.3888 0.0122 0.3834 2.1738
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.19478 0.17023 7.02 7.8e-12 ***
## FirstAuthorFemale1 -0.06244 0.06709 -0.93 0.35252
## LastAuthorFemale1 0.04314 0.06938 0.62 0.53434
## UniqueAuthors2 0.20936 0.06094 3.44 0.00064 ***
## UniqueAuthors3 0.18393 0.09458 1.94 0.05240 .
## UniqueAuthors4 0.20618 0.15751 1.31 0.19117
## UniqueAuthors5 0.73315 0.09329 7.86 2.6e-14 ***
## Year1997 -0.00465 0.29608 -0.02 0.98748
## Year1998 0.02020 0.20512 0.10 0.92159
## Year1999 0.03357 0.19723 0.17 0.86491
```

```

## Year2000      -0.27599      0.21992      -1.25      0.21012
## Year2001      -0.29477      0.22891      -1.29      0.19848
## Year2002       0.13764      0.22786       0.60      0.54611
## Year2003      -0.08408      0.21892      -0.38      0.70109
## Year2004      -0.25858      0.23688      -1.09      0.27556
## Year2005      -0.26177      0.20310      -1.29      0.19806
## Year2006      -0.16857      0.20054      -0.84      0.40101
## Year2007      -0.39460      0.19590      -2.01      0.04455 *
## Year2008       0.00108      0.18267       0.01      0.99527
## Year2009      -0.06171      0.18801      -0.33      0.74288
## Year2010      -0.19933      0.19452      -1.02      0.30601
## Year2011      -0.07809      0.19942      -0.39      0.69554
## Year2012      -0.24557      0.19534      -1.26      0.20931
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.592
## Multiple R-squared:  0.0881, Adjusted R-squared:  0.0455
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 453 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.148  0.869  0.952   0.912   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      2.02e-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 1.551 1      1.246
## LastAuthorFemale  1.624 1      1.274
## Year              1.206 16      1.006

```

Residuals from first and last author



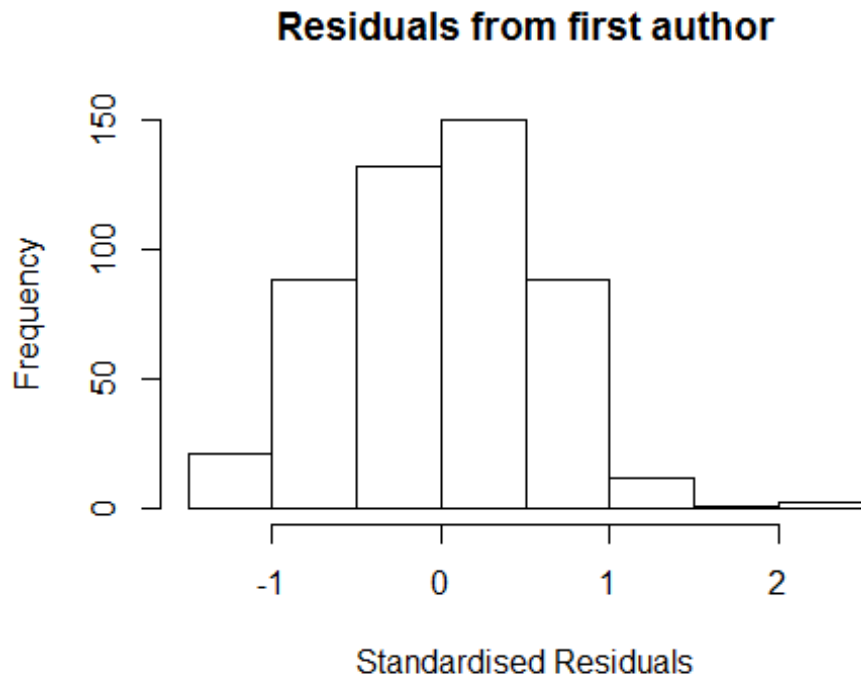
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3093 -0.4077 0.0219 0.4137 2.0893
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2690 0.1650 7.69 8.6e-14 ***
## FirstAuthorFemale1 -0.0234 0.0679 -0.34 0.731
## LastAuthorFemale1 0.0316 0.0715 0.44 0.659
## Year1997 -0.0141 0.2895 -0.05 0.961
## Year1998 0.0204 0.1989 0.10 0.918
## Year1999 0.0404 0.1926 0.21 0.834
## Year2000 -0.2626 0.2117 -1.24 0.216
## Year2001 -0.3032 0.2382 -1.27 0.204
## Year2002 0.1618 0.2267 0.71 0.476
## Year2003 -0.0931 0.2182 -0.43 0.670
## Year2004 -0.2773 0.2389 -1.16 0.246
## Year2005 -0.2578 0.2003 -1.29 0.199
```



```

## Year2006          -0.1582      0.2004   -0.79    0.430
## Year2007          -0.3523      0.1911   -1.84    0.066 .
## Year2008           0.0135      0.1774    0.08    0.939
## Year2009          -0.0175      0.1848   -0.09    0.924
## Year2010          -0.1599      0.1893   -0.84    0.399
## Year2011          -0.0245      0.1939   -0.13    0.900
## Year2012          -0.2046      0.1916   -1.07    0.286
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.612
## Multiple R-squared:  0.0522, Adjusted R-squared:  0.0163
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 44 weights are ~= 1. The remaining 450 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.220  0.875   0.950   0.914   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      2.02e-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 1.073 1          1.036
## Year              1.073 16          1.002

```

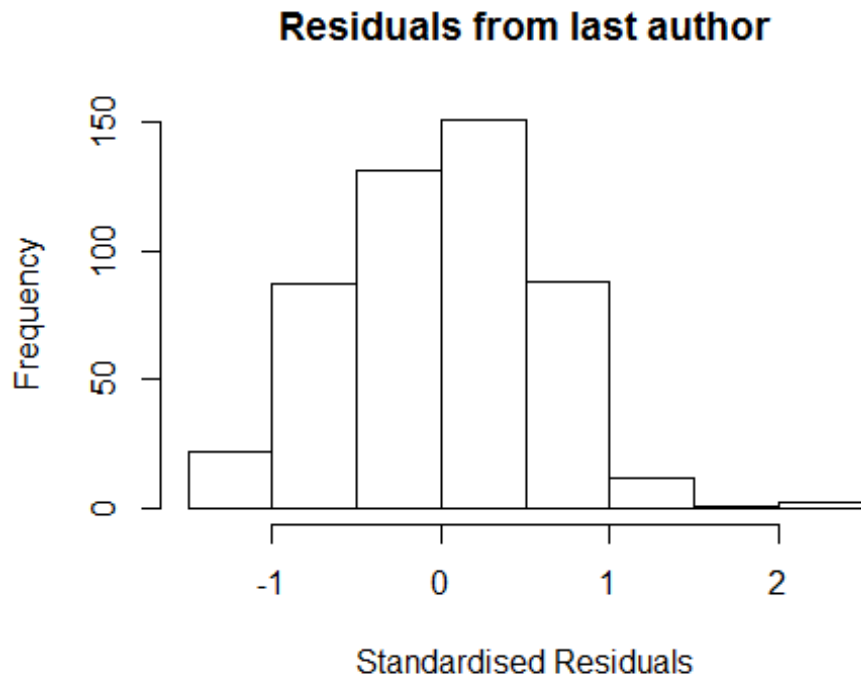


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3121 -0.4188  0.0163  0.4160  2.0859
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.27488    0.16715   7.63 1.3e-13 ***
## FirstAuthorFemale1 -0.00575    0.05660  -0.10  0.919
## Year1997        -0.01889    0.28955  -0.07  0.948
## Year1998         0.02061    0.19931   0.10  0.918
## Year1999         0.03718    0.19393   0.19  0.848
## Year2000        -0.26481    0.21453  -1.23  0.218
## Year2001        -0.30662    0.24025  -1.28  0.202
## Year2002         0.15842    0.22820   0.69  0.488
## Year2003        -0.09050    0.22050  -0.41  0.682
## Year2004        -0.27583    0.24118  -1.14  0.253
## Year2005        -0.25771    0.20106  -1.28  0.201
## Year2006        -0.16079    0.20110  -0.80  0.424
```

```

## Year2007      -0.35650    0.19276   -1.85    0.065 .
## Year2008      0.01002    0.17923    0.06    0.955
## Year2009     -0.02331    0.18661   -0.12    0.901
## Year2010     -0.15958    0.18999   -0.84    0.401
## Year2011     -0.02427    0.19524   -0.12    0.901
## Year2012     -0.20153    0.19214   -1.05    0.295
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.599
## Multiple R-squared:  0.0522, Adjusted R-squared:  0.0183
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 48 weights are ~= 1. The remaining 446 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.200  0.868  0.946  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      2.02e-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 1.12 1      1.058
## Year              1.12 16      1.004

```



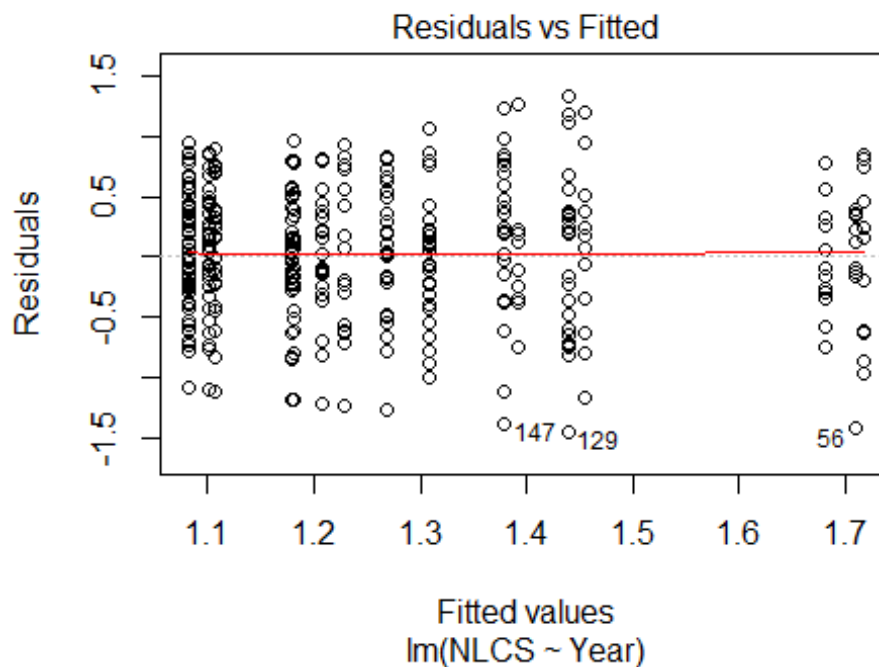
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3059 -0.4191 0.0213 0.4083 2.0969
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.26869 0.16604 7.64 1.2e-13 ***
## LastAuthorFemale1 0.01781 0.05963 0.30 0.765
## Year1997 -0.01834 0.28971 -0.06 0.950
## Year1998 0.01937 0.20015 0.10 0.923
## Year1999 0.03681 0.19381 0.19 0.849
## Year2000 -0.26682 0.21400 -1.25 0.213
## Year2001 -0.30778 0.23973 -1.28 0.200
## Year2002 0.15674 0.22886 0.68 0.494
## Year2003 -0.09614 0.22087 -0.44 0.664
## Year2004 -0.27922 0.24106 -1.16 0.247
## Year2005 -0.26046 0.20172 -1.29 0.197
## Year2006 -0.16556 0.20186 -0.82 0.413
```

```

## Year2007          -0.35780      0.19186      -1.86      0.063 .
## Year2008           0.00872      0.17866       0.05      0.961
## Year2009          -0.02520      0.18549      -0.14      0.892
## Year2010          -0.16346      0.19088      -0.86      0.392
## Year2011          -0.02907      0.19496      -0.15      0.882
## Year2012          -0.20698      0.19305      -1.07      0.284
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.599
## Multiple R-squared:  0.0524, Adjusted R-squared:  0.0185
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 43 weights are ~= 1. The remaining 451 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.194  0.868  0.948  0.910  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      2.02e-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: 494"
## [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
##   13   18   15   14   17   29   35   23   23   34   22   32   30   33   35
## 2011 2012
##   59   40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11   15   10    9   15   23   26   16   21   32   22   30   28   29   27
## 2011 2012

```

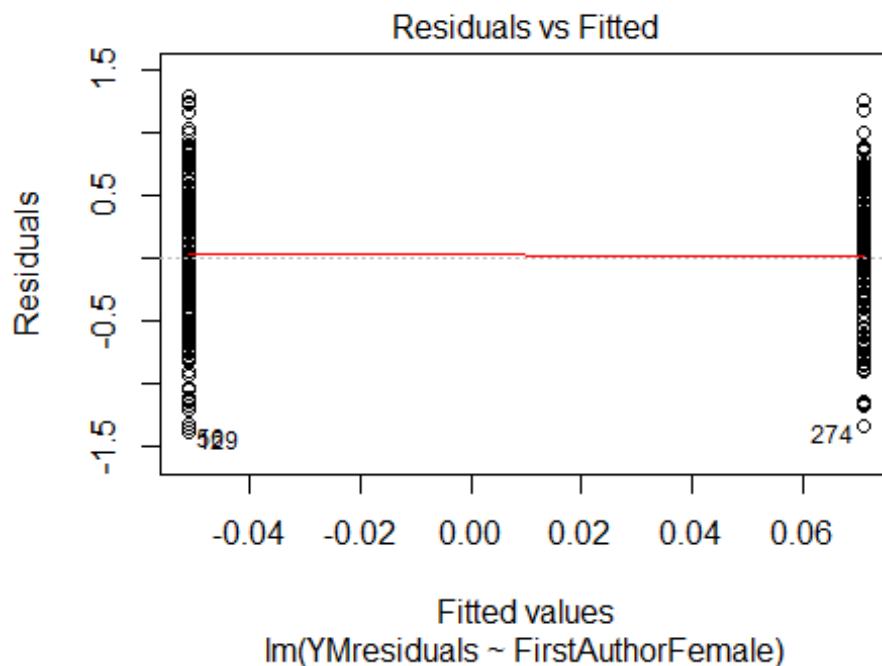
```
## 45 35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 15 9 9 14 22 24 15 18 31 19 28 25 28 27
## 2011 2012
## 44 33
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.4, df = 1, p-value = 0.06
## [1] "Female first author team size 2018 geometric mean: 1.69038696787705"
## [1] "Male first author team size 2018 geometric mean: 1.99133283418423"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

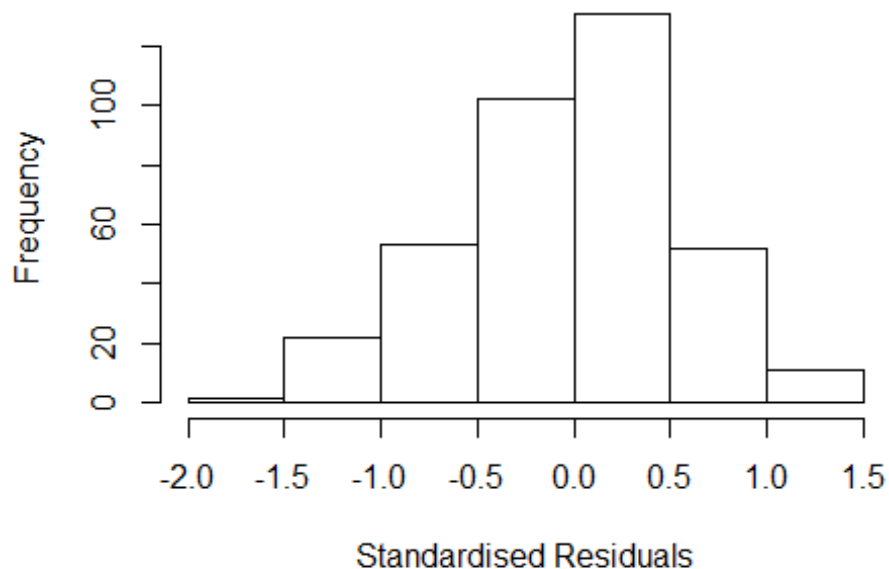
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 210, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.80956617726493"
## [1] "Male last author team size 2018 geometric mean: 1.88326122098113"

## 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 = 0.7
## 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 1.772  1      1.331
## LastAuthorFemale  1.983  1      1.408
## UniqueAuthors    2.296  4      1.109
## Year              2.178 16      1.025
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.5551 -0.3789 0.0378 0.3634 1.2754
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6360 0.2230 7.34 1.5e-12 ***
## FirstAuthorFemale1 0.1335 0.0745 1.79 0.0739 .
## LastAuthorFemale1 0.0216 0.0725 0.30 0.7664
## UniqueAuthors2 0.1949 0.0651 3.00 0.0029 **
## UniqueAuthors3 -0.0632 0.1143 -0.55 0.5805
## UniqueAuthors4 -0.1066 0.1032 -1.03 0.3026
## UniqueAuthors5 -0.2145 1.1444 -0.19 0.8515
## Year1997 -0.2574 0.3201 -0.80 0.4218
## Year1998 0.1094 0.2584 0.42 0.6722
## Year1999 -0.4141 0.2676 -1.55 0.1226
```

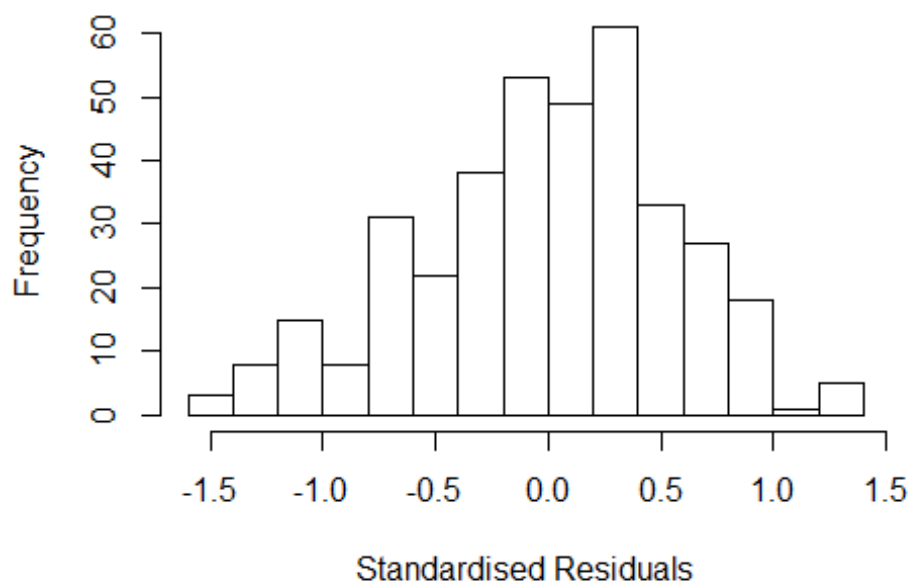


```

## Year2000          -0.0633      0.2436   -0.26   0.7953
## Year2001          -0.3648      0.2840   -1.28   0.1997
## Year2002          -0.2974      0.2929   -1.02   0.3107
## Year2003          -0.4512      0.2801   -1.61   0.1080
## Year2004          -0.5704      0.2478   -2.30   0.0219 *
## Year2005          -0.4308      0.2418   -1.78   0.0757 .
## Year2006          -0.6205      0.2517   -2.46   0.0142 *
## Year2007          -0.4360      0.2392   -1.82   0.0692 .
## Year2008          -0.5805      0.2449   -2.37   0.0183 *
## Year2009          -0.6946      0.2300   -3.02   0.0027 **
## Year2010          -0.5999      0.2516   -2.38   0.0177 *
## Year2011          -0.6415      0.2354   -2.72   0.0068 **
## Year2012          -0.6196      0.2388   -2.60   0.0099 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.568
## Multiple R-squared:  0.132, Adjusted R-squared:  0.0777
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 342 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.434  0.878  0.955  0.911  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          2.69e-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 1.616 1          1.271
## LastAuthorFemale  1.575 1          1.255
## Year              1.424 16          1.011

```

Residuals from first and last author



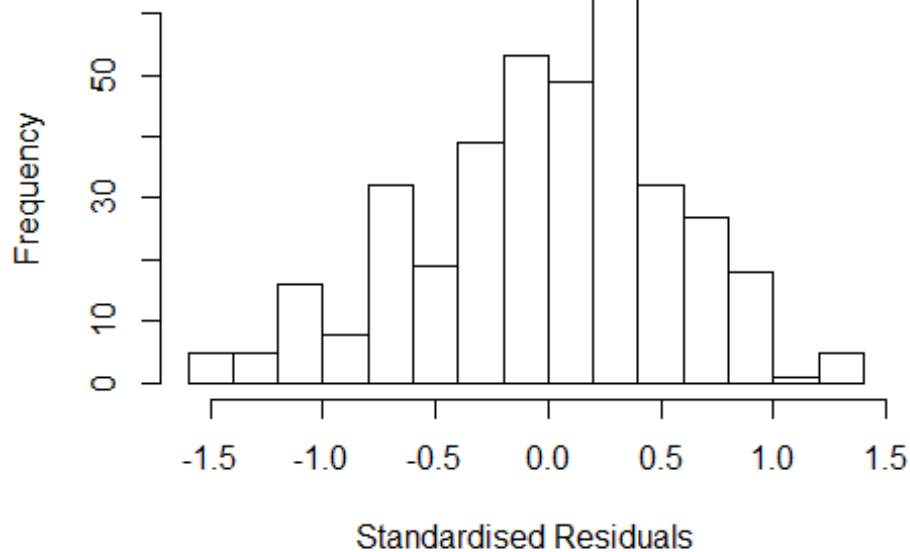
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4536 -0.3592 0.0299 0.3497 1.3134
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7218 0.2412 7.14 5.4e-12 ***
## FirstAuthorFemale1 0.1375 0.0726 1.89 0.0590 .
## LastAuthorFemale1 0.0331 0.0709 0.47 0.6410
## Year1997 -0.2919 0.3305 -0.88 0.3776
## Year1998 0.0229 0.2739 0.08 0.9335
## Year1999 -0.4283 0.2865 -1.49 0.1358
## Year2000 -0.0970 0.2672 -0.36 0.7167
## Year2001 -0.4096 0.3059 -1.34 0.1815
## Year2002 -0.3257 0.3087 -1.06 0.2921
## Year2003 -0.4921 0.3024 -1.63 0.1046
## Year2004 -0.5677 0.2673 -2.12 0.0344 *
## Year2005 -0.4673 0.2626 -1.78 0.0760 .
```

```

## Year2006          -0.6280      0.2739   -2.29   0.0224 *
## Year2007          -0.4606      0.2607   -1.77   0.0782 .
## Year2008          -0.6170      0.2619   -2.36   0.0190 *
## Year2009          -0.7456      0.2552   -2.92   0.0037 **
## Year2010          -0.6515      0.2731   -2.39   0.0176 *
## Year2011          -0.6817      0.2578   -2.64   0.0085 **
## Year2012          -0.6690      0.2605   -2.57   0.0106 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.571
## Multiple R-squared:  0.107, Adjusted R-squared:  0.0611
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 339 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.497  0.869  0.955  0.908  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      2.69e-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 1.244 1          1.115
## Year              1.244 16          1.007

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.4542 -0.3564 0.0318 0.3561 1.2940
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7280 0.2449 7.06 9e-12 ***
## FirstAuthorFemale1 0.1546 0.0638 2.42 0.0159 *
## Year1997 -0.2948 0.3336 -0.88 0.3775
## Year1998 0.0172 0.2769 0.06 0.9505
## Year1999 -0.4262 0.2901 -1.47 0.1426
## Year2000 -0.0937 0.2713 -0.35 0.7299
## Year2001 -0.4040 0.3081 -1.31 0.1906
## Year2002 -0.3256 0.3112 -1.05 0.2961
## Year2003 -0.4943 0.3049 -1.62 0.1059
## Year2004 -0.5650 0.2720 -2.08 0.0385 *
## Year2005 -0.4704 0.2662 -1.77 0.0781 .
## Year2006 -0.6296 0.2774 -2.27 0.0238 *
```

```

## Year2007          -0.4593      0.2648   -1.73   0.0837 .
## Year2008          -0.6174      0.2659   -2.32   0.0208 *
## Year2009          -0.7425      0.2588   -2.87   0.0044 **
## Year2010          -0.6499      0.2764   -2.35   0.0193 *
## Year2011          -0.6838      0.2614   -2.62   0.0093 **
## Year2012          -0.6738      0.2642   -2.55   0.0112 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.106, Adjusted R-squared:  0.0633
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 336 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.493  0.866  0.955  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      2.69e-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 0 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
## -1.5367 -0.3504  0.0178  0.3579  1.4290

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7081    0.2344    7.29 2.1e-12 ***
## LastAuthorFemale1  0.1039    0.0626    1.66  0.0980 .
## Year1997         -0.2648    0.3231   -0.82  0.4130
## Year1998          0.0369    0.2678    0.14  0.8906
## Year1999         -0.4074    0.2820   -1.44  0.1494
## Year2000         -0.0624    0.2601   -0.24  0.8106
## Year2001         -0.3742    0.2998   -1.25  0.2129
## Year2002         -0.2754    0.3033   -0.91  0.3645
## Year2003         -0.4494    0.3015   -1.49  0.1370
## Year2004         -0.5509    0.2604   -2.12  0.0351 *
## Year2005         -0.4169    0.2518   -1.66  0.0986 .
## Year2006         -0.5620    0.2640   -2.13  0.0339 *
## Year2007         -0.4272    0.2553   -1.67  0.0952 .
## Year2008         -0.5604    0.2527   -2.22  0.0272 *
## Year2009         -0.6933    0.2480   -2.80  0.0055 **
## Year2010         -0.6150    0.2642   -2.33  0.0205 *
## Year2011         -0.6145    0.2483   -2.47  0.0138 *
## Year2012         -0.6131    0.2538   -2.42  0.0162 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.57
## Multiple R-squared:  0.0988, Adjusted R-squared:  0.0556
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 340 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.447  0.866  0.956  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      2.69e-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: 372"
## [1] ""

```

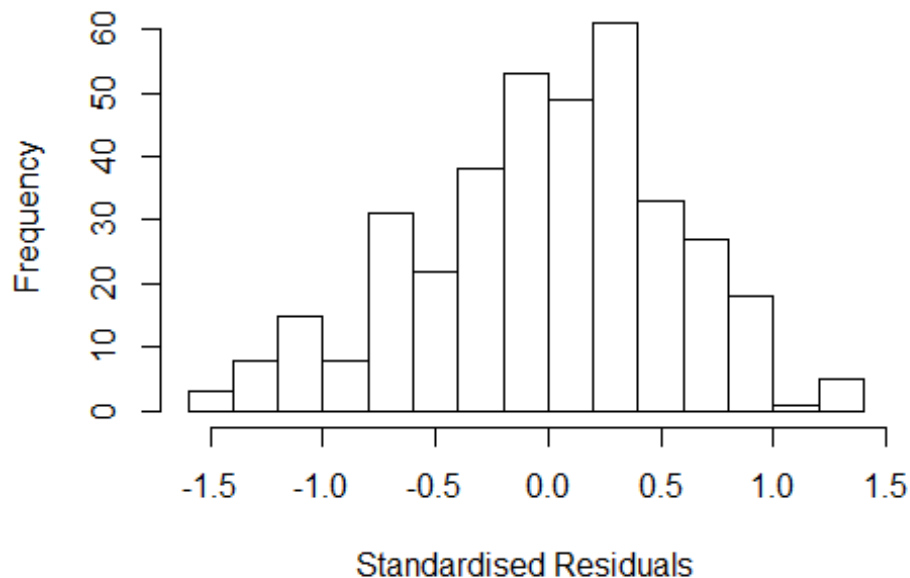
```

## [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
##    1    5    2    3    3    1    4    3    1    3    3    4    5    5    5
## 2011 2012
##    3    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    5    2    3    3    0    4    3    1    2    3    3    5    5    4
## 2011 2012
##    3    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    5    1    3    3    0    3    3    1    2    3    3    5    4    4
## 2011 2012
##    2    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 1"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 1.31607401295249"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.597e-06 1      0.001264
## LastAuthorFemale  2.449e+03 1      49.487620
## UniqueAuthors    -1.190e+18 3         NaN
## Year              5.385e+21 15      5.301189
##
## [1] "List of 0 outliers with 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
## -1.28e+00 -2.05e-01 -1.11e-16  2.11e-01  5.48e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.3840      0.2855   1.35  0.19022
## FirstAuthorFemale1 0.0127      0.2624   0.05  0.96177
## LastAuthorFemale1  0.0559      0.3297   0.17  0.86659
## UniqueAuthors2     0.4753      0.1586   3.00  0.00592 **
## UniqueAuthors3     0.2879      0.2557   1.13  0.27062
## UniqueAuthors4    -0.0822      0.6568  -0.13  0.90139
## Year1997           0.5807      0.2309   2.52  0.01840 *
## Year1998           0.7043      0.1586   4.44  0.00015 ***
## Year1999           0.6916      0.3156   2.19  0.03757 *
## Year2000           0.2561      0.2672   0.96  0.34665
## Year2002          -0.2560      0.0881  -2.91  0.00738 **
## Year2003           0.3212      0.3663   0.88  0.38867
## Year2004           0.5493      0.1586   3.46  0.00186 **
## Year2005           0.5443      0.2485   2.19  0.03769 *
## Year2006           0.2340      0.2719   0.86  0.39722
## Year2007           0.8973      0.9438   0.95  0.35048
## Year2008           0.1104      0.3243   0.34  0.73626
## Year2009           1.0092      0.3015   3.35  0.00249 **
## Year2010           0.0661      0.4165   0.16  0.87505
## Year2011           0.6555      0.2876   2.28  0.03110 *
## Year2012           0.5836      0.2361   2.47  0.02031 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.49,   Adjusted R-squared:  0.0968
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 38 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.338  0.899  0.949  0.923  0.983  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          2.13e-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

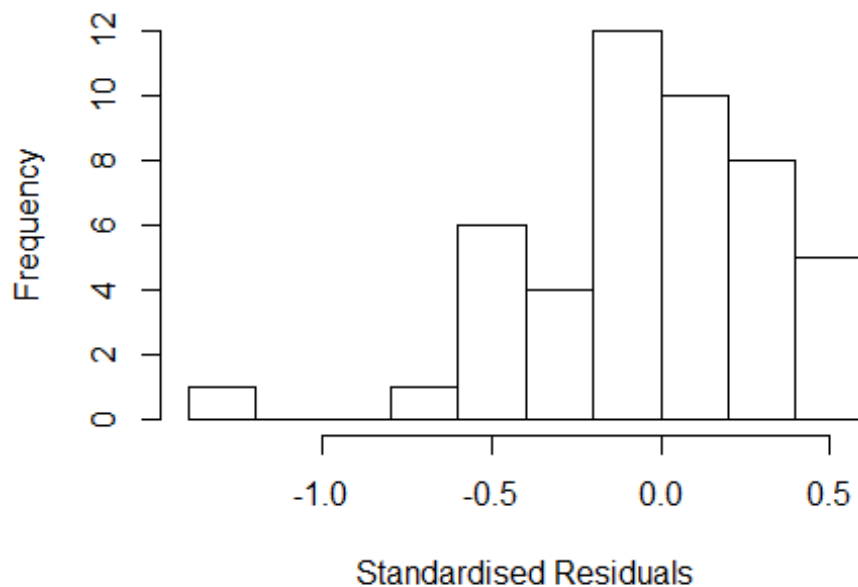
```

```
##               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

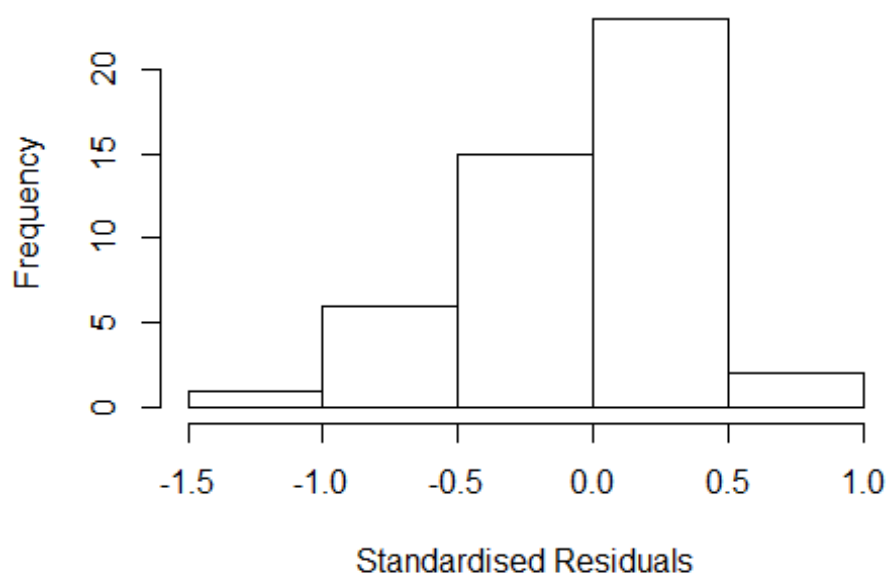
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author and team size



```
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## LastAuthorFemale  NaN 1             NaN
## Year              NaN 15            NaN
```

Residuals from first and last author



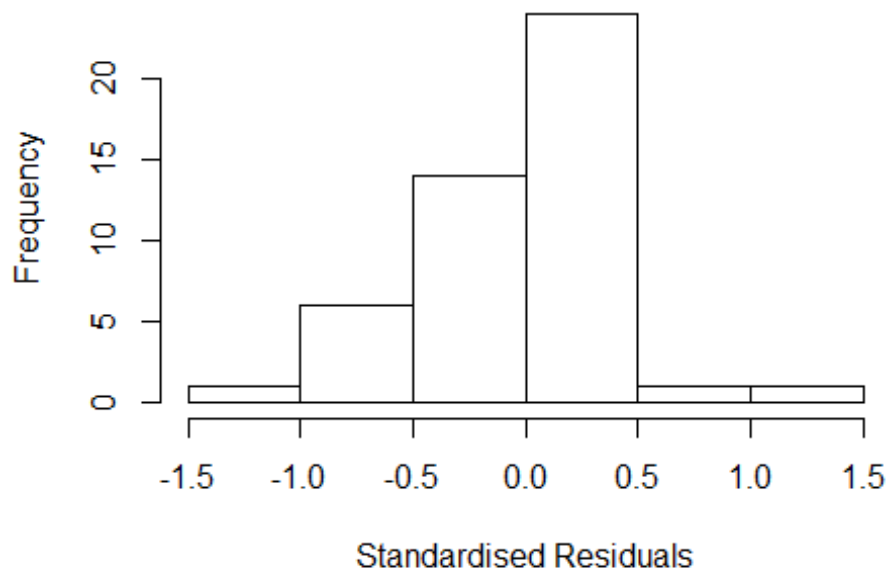
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2015 -0.2445 0.0612 0.2418 0.9542
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8336 0.2351 3.55 0.0014 **
## FirstAuthorFemale1 -0.0136 0.2775 -0.05 0.9612
## LastAuthorFemale1 0.1080 0.3265 0.33 0.7432
## Year1997 0.2252 0.2498 0.90 0.3747
## Year1998 0.2290 0.0000 Inf <2e-16 ***
## Year1999 0.4830 0.2108 2.29 0.0294 *
## Year2000 -0.0346 0.2457 -0.14 0.8889
## Year2002 -0.3094 0.1023 -3.02 0.0052 **
## Year2003 -0.1356 0.3041 -0.45 0.6590
## Year2004 0.0740 0.0000 Inf <2e-16 ***
## Year2005 0.3324 0.2494 1.33 0.1931
## Year2006 -0.1240 0.3146 -0.39 0.6962
```

```

## Year2007          0.3679      0.4931      0.75      0.4616
## Year2008         -0.3218      0.4200     -0.77      0.4498
## Year2009          0.6190      0.2301      2.69      0.0117 *
## Year2010         -0.3336      0.6167     -0.54      0.5927
## Year2011          0.4567      0.3097      1.47      0.1511
## Year2012          0.3192      0.1999      1.60      0.1211
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.398, Adjusted R-squared:  0.0454
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.399  0.893  0.968  0.915  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      2.13e-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 -7374 1      NaN
## Year              -7374 15      NaN

```

Residuals from first author



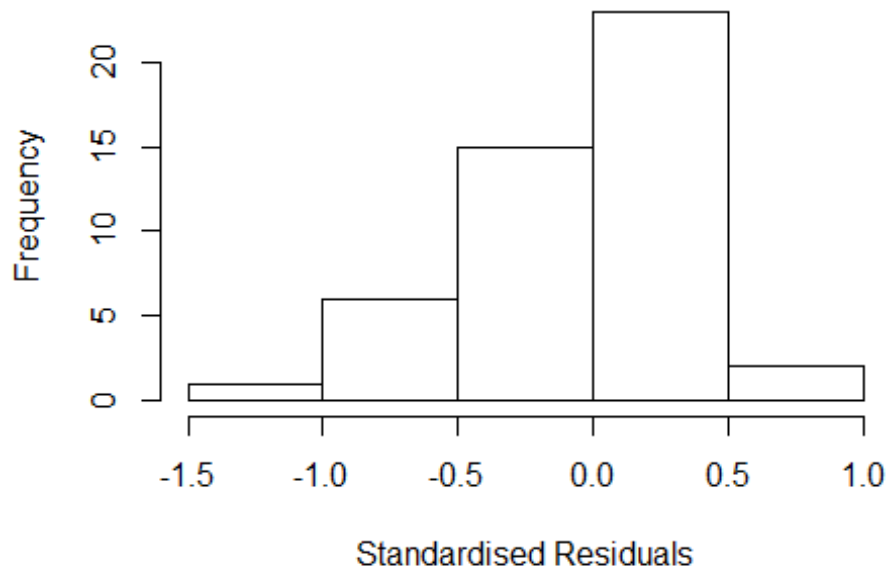
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1985 -0.2057 0.0483 0.2303 1.0448
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.63e-01 1.93e-01 4.47e+00 0.0001 ***
## FirstAuthorFemale1 6.51e-02 1.93e-01 3.40e-01 0.7383
## Year1997 2.09e-01 2.42e-01 8.60e-01 0.3954
## Year1998 2.29e-01 7.57e-09 3.02e+07 <2e-16 ***
## Year1999 4.61e-01 1.49e-01 3.09e+00 0.0043 **
## Year2000 -6.39e-02 2.06e-01 -3.10e-01 0.7585
## Year2002 -3.19e-01 8.85e-02 -3.61e+00 0.0011 **
## Year2003 -1.56e-01 2.88e-01 -5.40e-01 0.5920
## Year2004 7.40e-02 1.74e-08 4.26e+06 <2e-16 ***
## Year2005 3.03e-01 2.10e-01 1.44e+00 0.1595
## Year2006 -1.46e-01 2.81e-01 -5.20e-01 0.6061
## Year2007 3.36e-01 5.26e-01 6.40e-01 0.5280
```

```

## Year2008          -3.34e-01   3.99e-01 -8.40e-01   0.4096
## Year2009          5.97e-01   2.01e-01  2.96e+00   0.0059 **
## Year2010         -3.58e-01   5.72e-01 -6.20e-01   0.5367
## Year2011          3.88e-01   1.80e-01  2.16e+00   0.0392 *
## Year2012          2.63e-01   1.73e-01  1.53e+00   0.1376
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.396, Adjusted R-squared:  0.0733
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.428  0.895  0.969  0.919  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      2.13e-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 2676 1          51.733
## Year             2676 15          1.301

```

Residuals from last author



```
## [1] "List of 0 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
## -1.230 -0.242 0.056 0.238 0.974
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.38e-01 2.42e-01 3.46e+00 0.0017 **
## LastAuthorFemale1 9.03e-02 2.42e-01 3.70e-01 0.7123
## Year1997 2.26e-01 2.49e-01 9.10e-01 0.3700
## Year1998 2.29e-01 2.21e-08 1.03e+07 <2e-16 ***
## Year1999 4.80e-01 2.15e-01 2.23e+00 0.0336 *
## Year2000 -3.87e-02 2.53e-01 -1.50e-01 0.8792
## Year2002 -3.11e-01 1.04e-01 -2.98e+00 0.0057 **
## Year2003 -1.39e-01 3.12e-01 -4.50e-01 0.6586
## Year2004 7.40e-02 2.36e-08 3.13e+06 <2e-16 ***
## Year2005 3.28e-01 2.56e-01 1.28e+00 0.2102
## Year2006 -1.26e-01 3.17e-01 -4.00e-01 0.6935
## Year2007 3.92e-01 4.69e-01 8.40e-01 0.4098
```

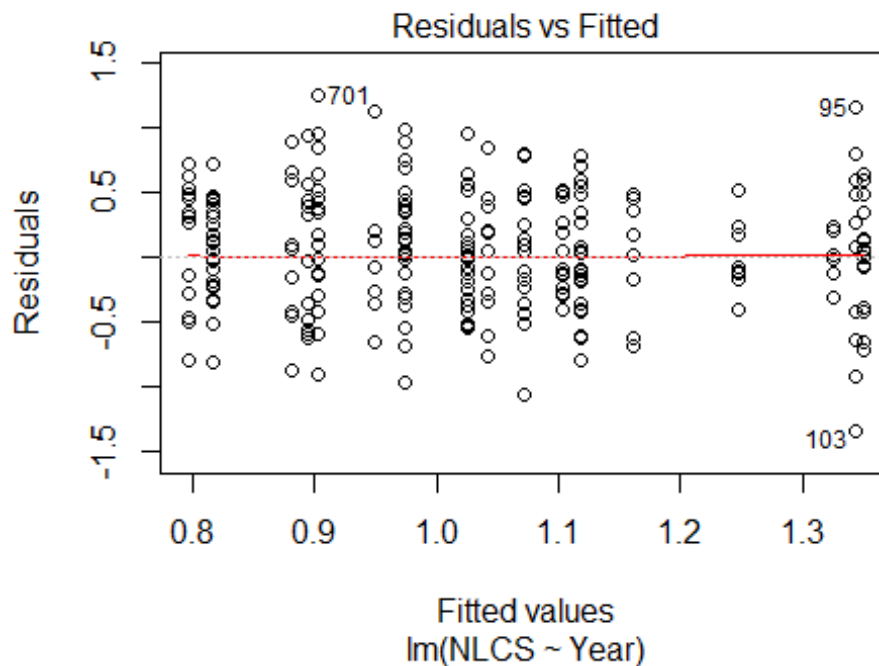
```

## Year2008          -3.28e-01   4.45e-01 -7.40e-01   0.4669
## Year2009           6.16e-01   2.38e-01  2.59e+00   0.0147 *
## Year2010          -3.44e-01   6.47e-01 -5.30e-01   0.5984
## Year2011           4.46e-01   3.02e-01  1.48e+00   0.1502
## Year2012           3.11e-01   2.06e-01  1.51e+00   0.1410
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.408, Adjusted R-squared:  0.0925
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.351  0.888  0.965  0.911  0.988  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      2.13e-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: 47"
## [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
##   33   49   44   46   37   42   42   28   37   31   22   51   45   40   42
## 2011 2012
##   42   40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   18   13   16   14    8    6    8    8   11   11   12   28   25   24   23
## 2011 2012
##   29   18

```



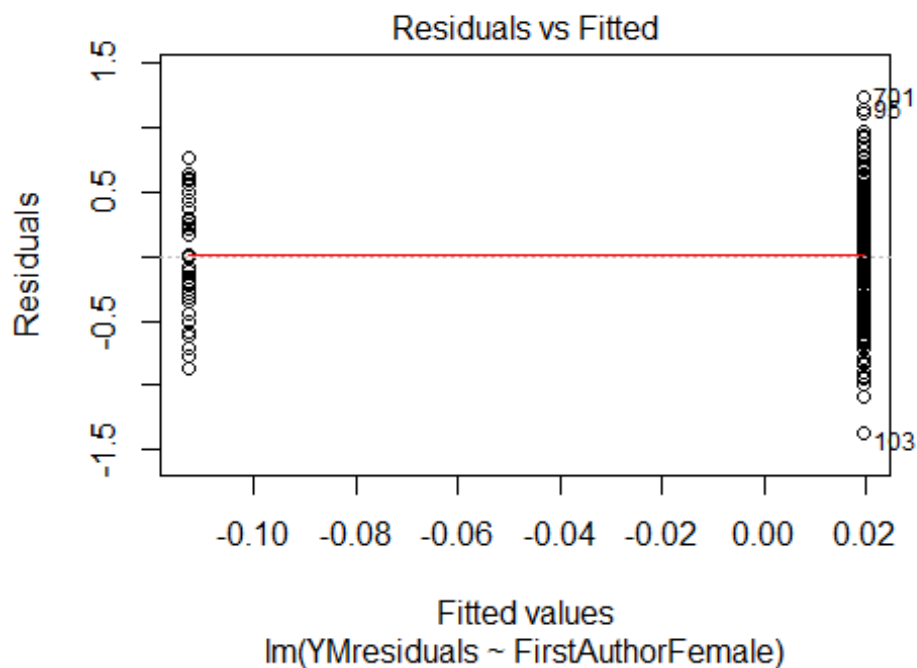
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 16 12 14 13 6 6 3 6 10 6 10 18 19 23 17
## 2011 2012
## 23 16
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 27, df = 16, p-value = 0.05
```



```
##
## 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: 4.11953428781424"
## [1] "Male first author team size 2018 geometric mean: 3.43291712546342"
## 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 = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male last author team size 2018 geometric mean: 3.57958765570427"

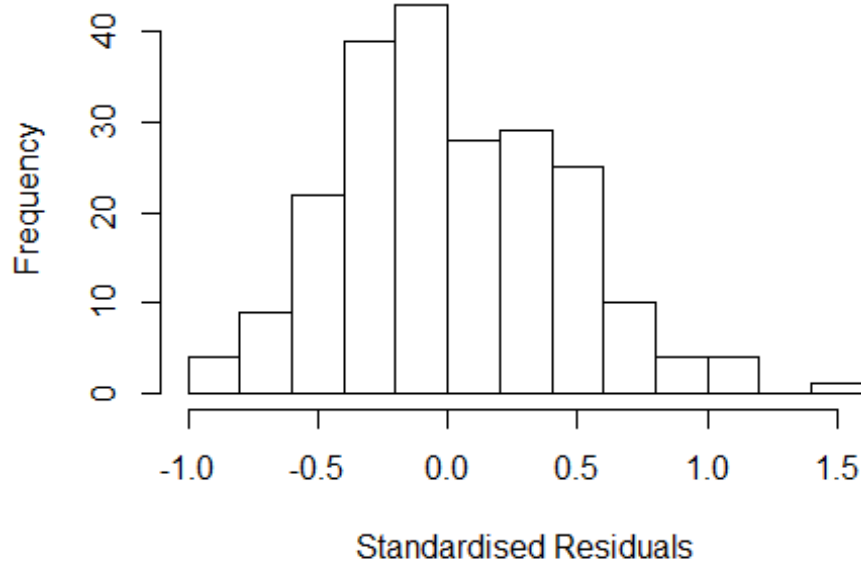
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16, 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	1.483	1	1.218
LastAuthorFemale	1.567	1	1.252
UniqueAuthors	3.355	4	1.163
Year	6.186	16	1.059

Residuals from first and last author and team size



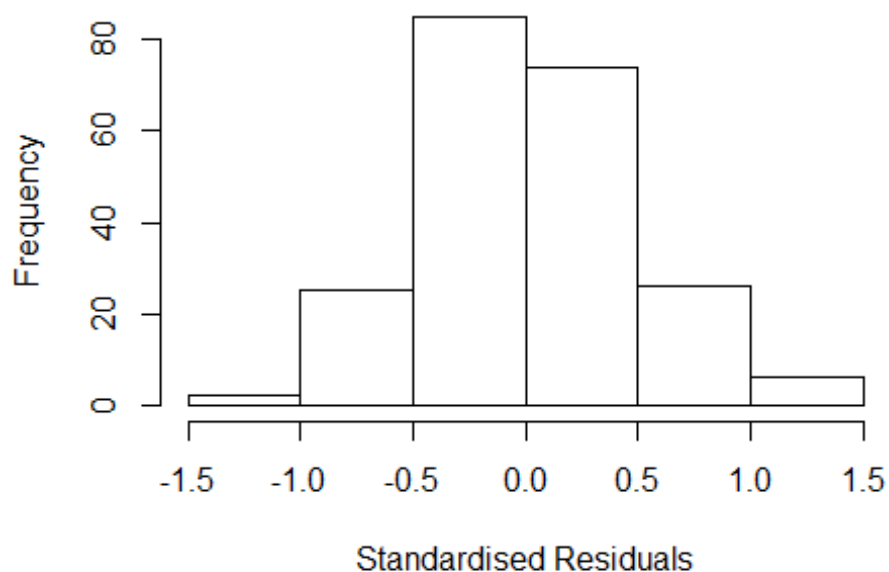
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9903 -0.2893 -0.0426 0.3325 1.5293
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8843 0.1608 5.50 1.2e-07 ***
## FirstAuthorFemale1 -0.0220 0.0794 -0.28 0.78212
## LastAuthorFemale1 -0.2895 0.0785 -3.69 0.00029 ***
## UniqueAuthors2 0.3226 0.1061 3.04 0.00269 **
## UniqueAuthors3 0.3825 0.1104 3.47 0.00065 ***
## UniqueAuthors4 0.4332 0.1099 3.94 0.00011 ***
## UniqueAuthors5 0.5542 0.1391 3.99 9.5e-05 ***
## Year1997 0.1060 0.2676 0.40 0.69239
## Year1998 0.1667 0.1783 0.93 0.35101
## Year1999 -0.0486 0.1849 -0.26 0.79292
```

```

## Year2000          -0.1690      0.2148   -0.79  0.43246
## Year2001           0.2504      0.2133    1.17  0.24180
## Year2002           0.0561      0.1988    0.28  0.77819
## Year2003          -0.3425      0.3220   -1.06  0.28884
## Year2004          -0.2143      0.2379   -0.90  0.36869
## Year2005           0.1280      0.2159    0.59  0.55388
## Year2006          -0.3919      0.2238   -1.75  0.08147 .
## Year2007          -0.2878      0.1798   -1.60  0.11095
## Year2008          -0.2243      0.1797   -1.25  0.21345
## Year2009          -0.0990      0.1730   -0.57  0.56785
## Year2010          -0.4153      0.1963   -2.12  0.03565 *
## Year2011          -0.3004      0.1782   -1.69  0.09353 .
## Year2012          -0.4693      0.2431   -1.93  0.05497 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.265, Adjusted R-squared:  0.183
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.219  0.884  0.952  0.913  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      4.59e-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 1.416 1      1.190
## LastAuthorFemale  1.495 1      1.223
## Year              2.024 16      1.022

```

Residuals from first and last author



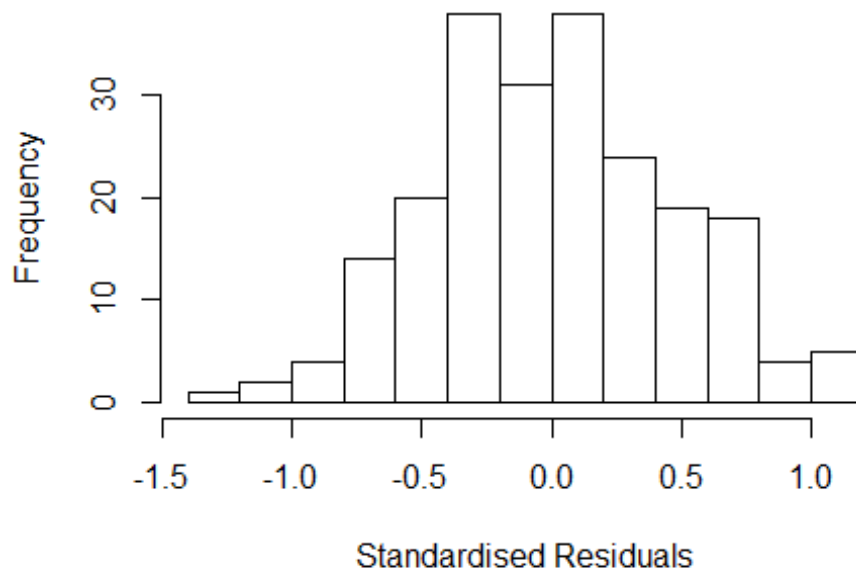
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3851 -0.2990 -0.0107 0.3245 1.1382
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1049 0.1385 7.98 1.1e-13 ***
## FirstAuthorFemale1 0.0110 0.0798 0.14 0.8904
## LastAuthorFemale1 -0.2159 0.0791 -2.73 0.0069 **
## Year1997 0.2802 0.2811 1.00 0.3200
## Year1998 0.2222 0.1645 1.35 0.1783
## Year1999 0.0334 0.1692 0.20 0.8437
## Year2000 -0.0846 0.2321 -0.36 0.7158
## Year2001 0.2531 0.1724 1.47 0.1437
## Year2002 0.1200 0.1735 0.69 0.4900
## Year2003 -0.1052 0.2725 -0.39 0.6999
## Year2004 -0.2767 0.2227 -1.24 0.2154
## Year2005 0.1436 0.2015 0.71 0.4770
```

```

## Year2006          -0.3493      0.2182    -1.60    0.1110
## Year2007          -0.1720      0.1952    -0.88    0.3794
## Year2008          -0.1044      0.1693    -0.62    0.5380
## Year2009          -0.0392      0.1678    -0.23    0.8154
## Year2010          -0.3851      0.2204    -1.75    0.0821 .
## Year2011          -0.2216      0.1622    -1.37    0.1734
## Year2012          -0.3303      0.2099    -1.57    0.1173
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.482
## Multiple R-squared:  0.173, Adjusted R-squared:  0.0986
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 196 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.390  0.880  0.949  0.915  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      4.59e-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 1.351 1          1.162
## Year              1.351 16          1.009

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.33016 -0.33109 -0.00252 0.29910 1.17684
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0890 0.1511 7.21 1.1e-11 ***
## FirstAuthorFemale1 -0.0326 0.0872 -0.37 0.709
## Year1997 0.2411 0.2932 0.82 0.412
## Year1998 0.1962 0.1813 1.08 0.281
## Year1999 -0.0044 0.1811 -0.02 0.981
## Year2000 -0.0628 0.2359 -0.27 0.790
## Year2001 0.2424 0.1685 1.44 0.152
## Year2002 0.1358 0.1837 0.74 0.461
## Year2003 -0.1407 0.2740 -0.51 0.608
## Year2004 -0.2911 0.2295 -1.27 0.206
## Year2005 0.1200 0.2183 0.55 0.583
## Year2006 -0.3682 0.2345 -1.57 0.118
```

```

## Year2007          -0.1700      0.2044   -0.83    0.406
## Year2008          -0.0884      0.1825   -0.48    0.628
## Year2009          -0.1060      0.1731   -0.61    0.541
## Year2010          -0.4028      0.2279   -1.77    0.079 .
## Year2011          -0.2589      0.1719   -1.51    0.134
## Year2012          -0.3485      0.2154   -1.62    0.107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.147, Adjusted R-squared:  0.0745
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 197 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.431  0.869  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      4.59e-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 1.484 1      1.218
## Year      1.484 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3858 -0.3016 -0.0074  0.3303  1.1394

```



```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1053    0.1388    7.96 1.2e-13 ***
## LastAuthorFemale1 -0.2140    0.0787   -2.72  0.0071 **
## Year1997          0.2806    0.2822    0.99  0.3214
## Year1998          0.2220    0.1648    1.35  0.1794
## Year1999          0.0359    0.1670    0.21  0.8301
## Year2000         -0.0835    0.2323   -0.36  0.7197
## Year2001          0.2541    0.1720    1.48  0.1411
## Year2002          0.1195    0.1737    0.69  0.4922
## Year2003         -0.1071    0.2728   -0.39  0.6949
## Year2004         -0.2759    0.2240   -1.23  0.2195
## Year2005          0.1426    0.2019    0.71  0.4806
## Year2006         -0.3500    0.2179   -1.61  0.1099
## Year2007         -0.1716    0.1955   -0.88  0.3813
## Year2008         -0.1019    0.1663   -0.61  0.5409
## Year2009         -0.0375    0.1680   -0.22  0.8238
## Year2010         -0.3866    0.2214   -1.75  0.0823 .
## Year2011         -0.2194    0.1602   -1.37  0.1724
## Year2012         -0.3280    0.2112   -1.55  0.1220
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.174, Adjusted R-squared:  0.104
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 197 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.380  0.882  0.952  0.913  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      4.59e-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: 218"
## [1] ""

```

```

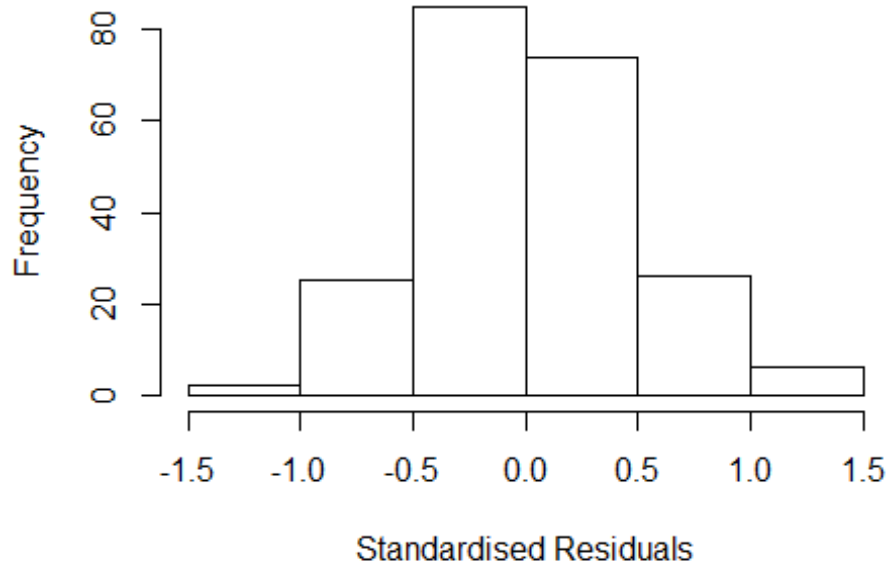
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1501"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 2000 2001 2002 2004 2005 2006 2007 2008 2010 2011
##    3    2    3    1    2    7    4    8    5    3    4    2    2
##
## 1996 1997 1998 2000 2001 2002 2004 2005 2006 2007 2008 2010 2011
##    1    1    0    1    0    1    0    6    2    3    4    2    1
##
## 1996 1997 1998 2000 2001 2002 2004 2005 2006 2007 2008 2010 2011
##    0    1    0    0    0    1    0    6    2    3    4    2    1
## [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: 4"
## [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 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

```

Residuals from last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## LastAuthorFemale  NaN 1             NaN
## Year              NaN 7             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 = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
## -9.28e-01 -9.02e-02  9.25e-16  1.11e-01  3.27e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.07e-15   0.00e+00  -Inf    < 2e-16 ***
## FirstAuthorFemale1  3.37e-01   1.03e-01   3.28  0.00833 **
## LastAuthorFemale1  -2.18e-02   1.29e-01  -0.17  0.86885
## Year2002          6.17e-01   0.00e+00   Inf    < 2e-16 ***
## Year2005          9.28e-01   1.85e-01   5.03  0.00052 ***
## Year2006          1.27e+00   8.26e-02  15.37  2.8e-08 ***
## Year2007          1.40e+00   2.28e-01   6.15  0.00011 ***
```

```

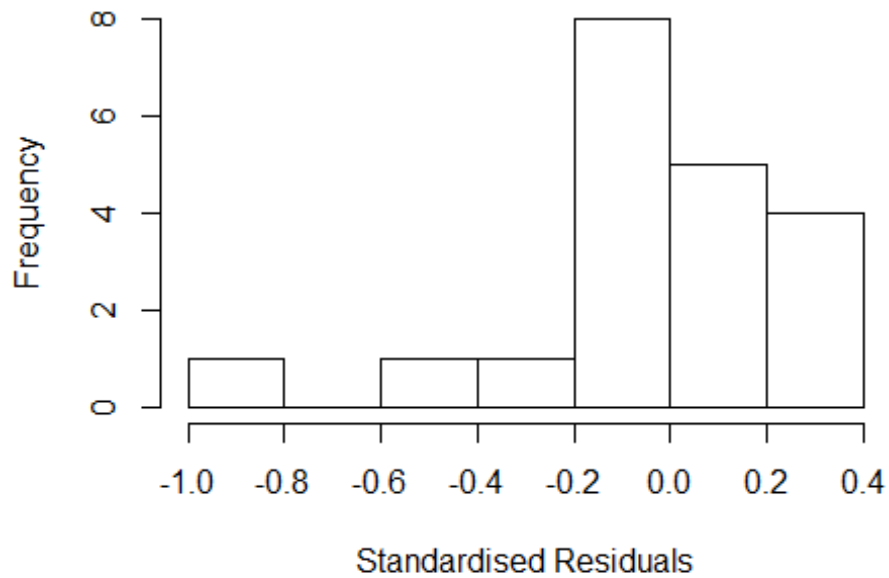
## Year2008          1.21e+00   8.26e-02   14.66   4.4e-08 ***
## Year2010          9.83e-01   1.29e-01    7.62   1.8e-05 ***
## Year2011          8.08e-01   1.03e-01    7.85   1.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.327
## Multiple R-squared:  0.762, Adjusted R-squared:  0.547
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 17 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.402  0.944  0.975  0.934  0.998  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      5.00e-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"

## 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

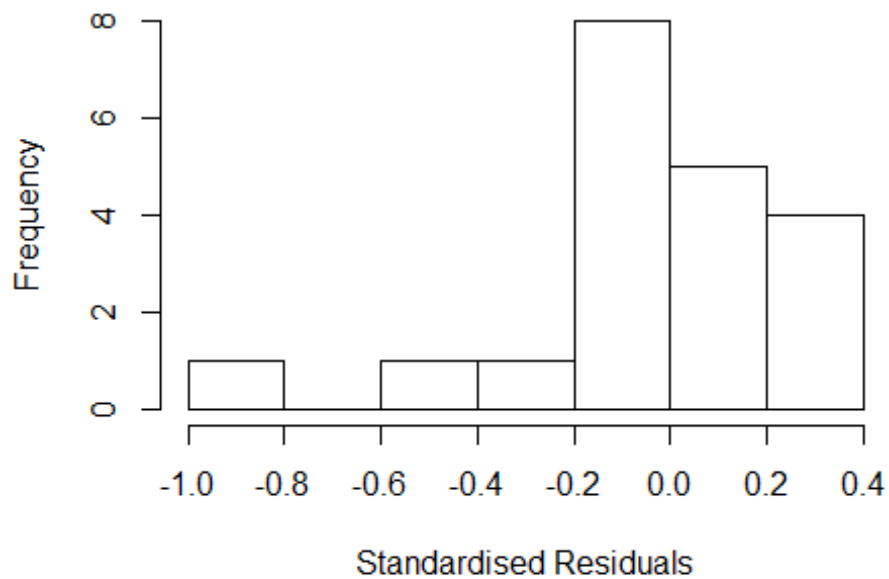
```

Residuals from first and last author



##		GVIF	Df	$GVIF^{(1/(2*Df))}$
##	FirstAuthorFemale	NaN	1	NaN
##	Year	NaN	7	NaN

Residuals from first author



```

## [1] "List of 0 outliers 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
## -9.41e-01 -9.00e-02 4.34e-17 1.10e-01 3.14e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -6.42e-16 0.00e+00 -Inf < 2e-16 ***
## FirstAuthorFemale1 3.37e-01 1.03e-01 3.26 0.0076 **
## Year2002 6.17e-01 0.00e+00 Inf < 2e-16 ***
## Year2005 9.41e-01 1.35e-01 6.95 2.4e-05 ***
## Year2006 1.27e+00 8.29e-02 15.33 9.1e-09 ***
## Year2007 1.41e+00 2.34e-01 6.00 8.9e-05 ***
## Year2008 1.21e+00 8.29e-02 14.61 1.5e-08 ***
## Year2010 9.72e-01 1.07e-01 9.13 1.8e-06 ***
## Year2011 8.08e-01 1.03e-01 7.81 8.2e-06 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.301
## Multiple R-squared: 0.777, Adjusted R-squared: 0.616
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 16 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.306 0.935 0.969 0.919 0.992 0.999
## Algorithmic parameters:
## tuning.chi bb tuning.psi refine.tol
## 1.55e+00 5.00e-01 4.69e+00 1.00e-07
## rel.tol solve.tol eps.outlier eps.x
## 1.00e-07 1.00e-07 5.00e-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 -122.8  1          NaN
## Year              -122.8  7          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
## -9.12e-01 -9.01e-02  4.04e-16  1.40e-01  3.43e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   -1.42e-16   1.54e-08   0.00e+00  1.00000
## LastAuthorFemale1 -1.13e-02   1.34e-01  -8.00e-02  0.93439
## Year2002         6.17e-01   1.54e-08  4.00e+07  < 2e-16 ***
## Year2005         9.12e-01   1.95e-01  4.68e+00  0.00067 ***
## Year2006         1.44e+00   5.60e-02  2.57e+01  3.6e-11 ***
## Year2007         1.74e+00   1.97e-01  8.80e+00  2.6e-06 ***
## Year2008         1.47e+00   1.39e-01  1.05e+01  4.5e-07 ***
## Year2010         1.32e+00   7.31e-02  1.80e+01  1.7e-09 ***
## Year2011         1.14e+00   1.54e-08  7.45e+07  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.718, Adjusted R-squared:  0.514
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 17 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.481  0.916  0.982  0.936  0.995  0.999
## Algorithmic parameters:
##           tuning.chi              bb           tuning.psi           refine.tol
##           1.55e+00           5.00e-01           4.69e+00           1.00e-07
##           rel.tol           solve.tol           eps.outlier           eps.x
##           1.00e-07           1.00e-07           5.00e-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: 20"
## [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
##    7   12   16    6   10    9    2   10    9    6   13   13   27   36   12
## 2011 2012
##   14    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    4    4    4    4    4    1    4    4    2    4    5   14   19    6
## 2011 2012
##    5    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    4    4    3    2    4    1    3    3    1    3    4   10   18    4
## 2011 2012
##    3    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.48480655243262"
## [1] "Male first author team size 2018 geometric mean: 3.95720489293585"

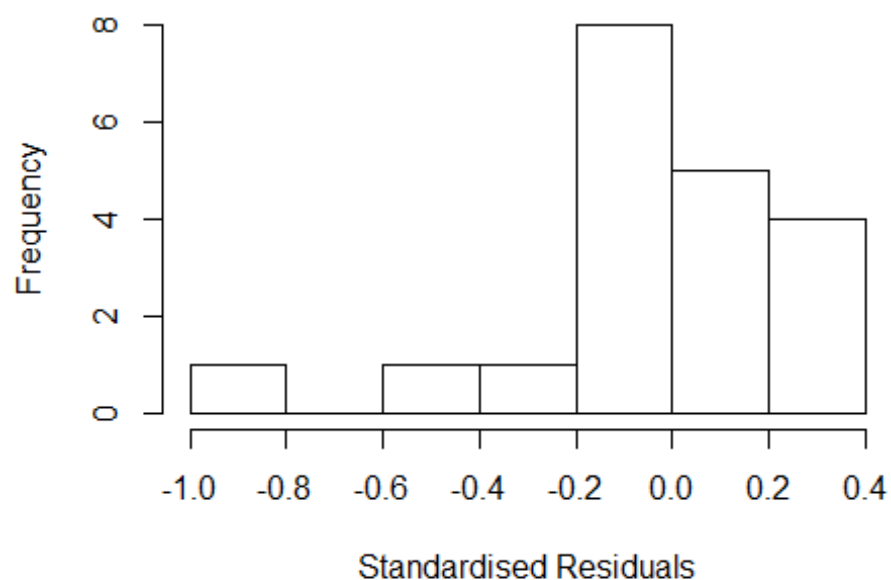
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 4.34365821992149"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

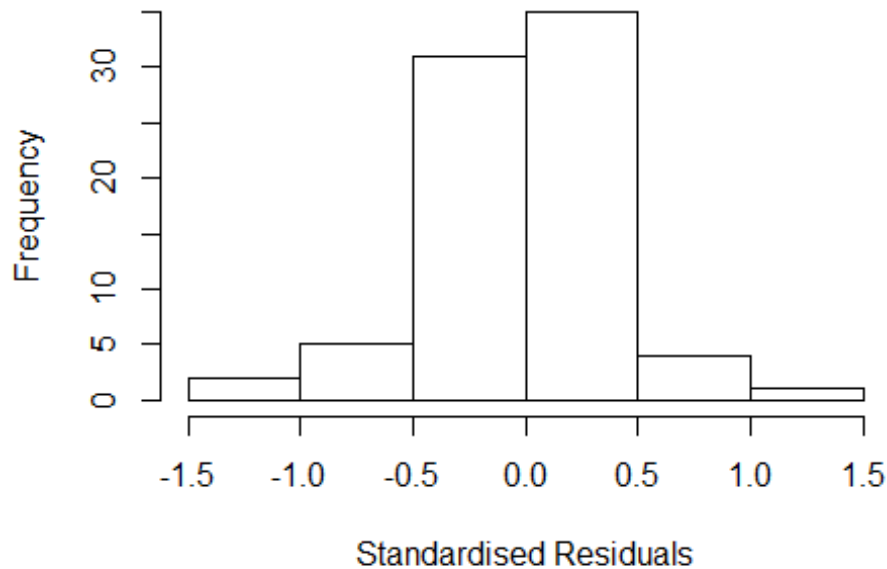
```


Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5.5, 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 3.522e+14  1      1.877e+07
## LastAuthorFemale  2.841e+00  1      1.686e+00
## UniqueAuthors    2.248e+18  4      1.968e+02
## Year              8.271e+18 16      3.901e+00
```

Residuals from first and last author and team size



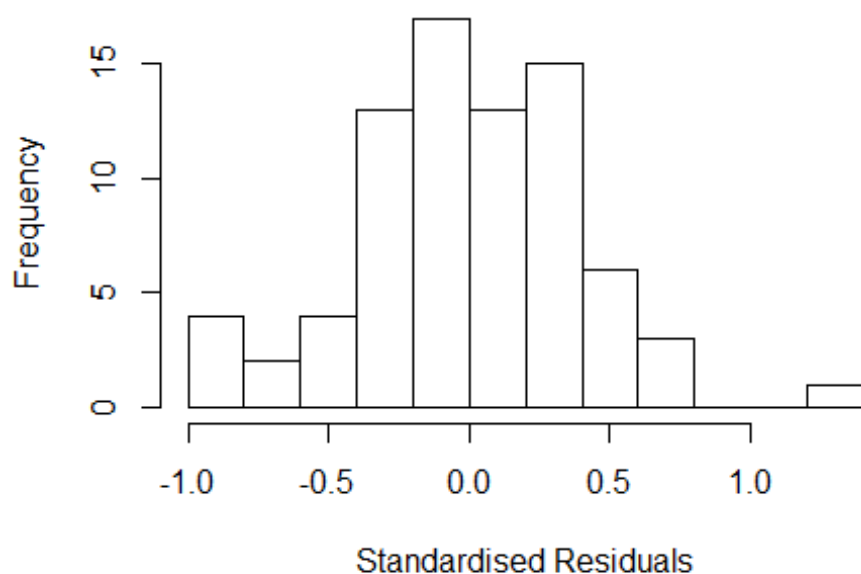
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0483 -0.2303 0.0125 0.2354 1.3445
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6264 0.1860 3.37 0.00139 **
## FirstAuthorFemale1 0.1416 0.1017 1.39 0.16966
## LastAuthorFemale1 -0.3907 0.1109 -3.52 0.00087 ***
## UniqueAuthors2 0.1951 0.2966 0.66 0.51341
## UniqueAuthors3 -0.1315 0.2834 -0.46 0.64437
## UniqueAuthors4 0.0329 0.3356 0.10 0.92231
## UniqueAuthors5 0.0432 0.3140 0.14 0.89097
## Year1997 0.5139 0.2559 2.01 0.04950 *
## Year1998 0.4080 0.3874 1.05 0.29694
## Year1999 0.2898 0.2570 1.13 0.26439
```

```

## Year2000          0.1168      0.3047      0.38  0.70301
## Year2001          0.2774      0.2394      1.16  0.25148
## Year2002          0.2796      0.1860      1.50  0.13847
## Year2003          0.6182      0.2218      2.79  0.00727 **
## Year2004          0.3319      0.2406      1.38  0.17337
## Year2005          0.1186      0.1980      0.60  0.55168
## Year2006          0.4942      0.2394      2.06  0.04368 *
## Year2007          0.1465      1.0006      0.15  0.88416
## Year2008          0.5152      0.2935      1.76  0.08474 .
## Year2009          0.2269      0.2411      0.94  0.35084
## Year2010          0.6653      0.2464      2.70  0.00920 **
## Year2011          0.1965      0.2580      0.76  0.44953
## Year2012          0.0777      0.1976      0.39  0.69553
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.294, Adjusted R-squared:  0.0121
## Convergence in 39 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.143  0.889  0.957  0.899  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          1.28e-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 1.775e+14 1      1.332e+07
## LastAuthorFemale  1.785e+00 1      1.336e+00
## Year              3.284e+14 16      2.842e+00

```

Residuals from first and last author



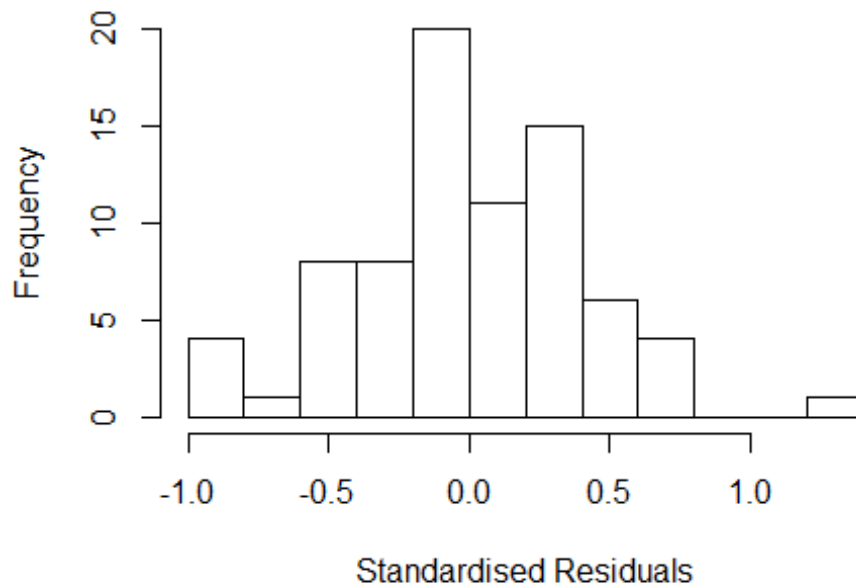
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.88e-01 -2.33e-01 -2.15e-16 2.31e-01 1.35e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6280 0.1484 4.23 8.2e-05 ***
## FirstAuthorFemale1 0.0826 0.1084 0.76 0.4493
## LastAuthorFemale1 -0.3845 0.1088 -3.53 0.0008 ***
## Year1997 0.4716 0.2222 2.12 0.0380 *
## Year1998 0.4381 0.3632 1.21 0.2325
## Year1999 0.4233 0.2093 2.02 0.0477 *
## Year2000 0.3397 0.2047 1.66 0.1023
## Year2001 0.3443 0.1984 1.74 0.0880 .
## Year2002 0.2780 0.1484 1.87 0.0659 .
## Year2003 0.6113 0.2871 2.13 0.0374 *
## Year2004 0.2516 0.1914 1.31 0.1938
## Year2005 0.0444 0.1437 0.31 0.7581
```

```

## Year2006          0.3610      0.1868      1.93      0.0581 .
## Year2007          0.2520      0.6767      0.37      0.7110
## Year2008          0.5098      0.1694      3.01      0.0038 **
## Year2009          0.2604      0.1972      1.32      0.1917
## Year2010          0.7215      0.1438      5.02      5.1e-06 ***
## Year2011          0.2535      0.1822      1.39      0.1693
## Year2012          0.1331      0.2143      0.62      0.5367
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.241, Adjusted R-squared:  0.00993
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.217  0.896  0.965  0.909  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      1.28e-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 9.237e+13 1      9.611e+06
## Year              9.237e+13 16      2.732e+00

```

Residuals from first author



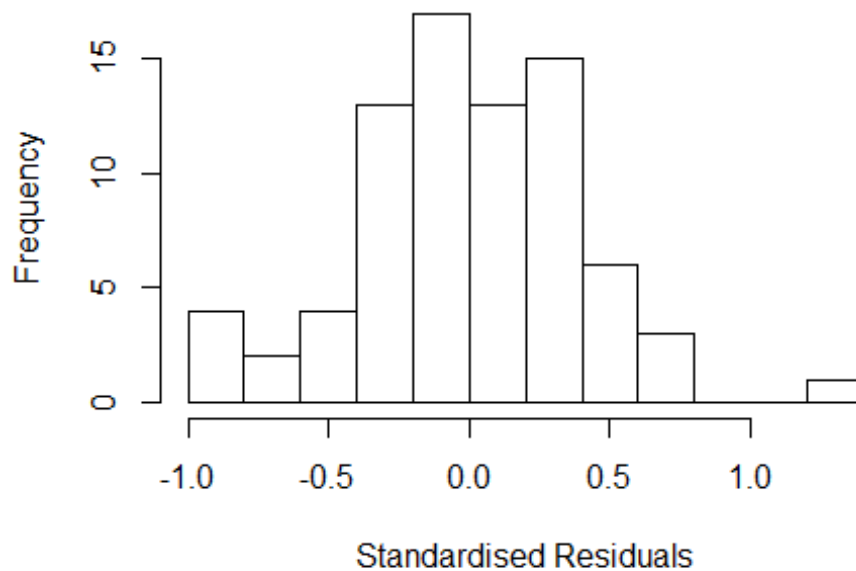
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.86855 -0.23275 -0.00686 0.25045 1.36549
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6380 0.1480 4.31 6.1e-05 ***
## FirstAuthorFemale1 0.0575 0.1103 0.52 0.604
## Year1997 0.3666 0.2437 1.50 0.138
## Year1998 0.4412 0.3375 1.31 0.196
## Year1999 0.2880 0.1694 1.70 0.094 .
## Year2000 0.3422 0.1969 1.74 0.087 .
## Year2001 0.1597 0.2761 0.58 0.565
## Year2002 0.2680 0.1480 1.81 0.075 .
## Year2003 0.6068 0.2863 2.12 0.038 *
## Year2004 0.1292 0.2855 0.45 0.653
## Year2005 0.0595 0.1463 0.41 0.686
## Year2006 0.3508 0.1859 1.89 0.064 .
```

```

## Year2007          0.2305      0.5625      0.41      0.683
## Year2008          0.4825      0.1828      2.64      0.011 *
## Year2009          0.2167      0.1870      1.16      0.251
## Year2010          0.7177      0.1453      4.94      6.6e-06 ***
## Year2011          0.2519      0.1866      1.35      0.182
## Year2012          0.0407      0.2020      0.20      0.841
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.177, Adjusted R-squared:  -0.0564
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.321  0.902  0.963  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      1.28e-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 2.047 1      1.431
## Year              2.047 16      1.023

```

Residuals from last author



```
## [1] "List of 0 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
## -9.06e-01 -2.30e-01 2.84e-16 2.34e-01 1.31e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6587 0.1471 4.48 3.4e-05 ***
## LastAuthorFemale1 -0.3744 0.1102 -3.40 0.0012 **
## Year1997 0.4638 0.2263 2.05 0.0448 *
## Year1998 0.4291 0.3320 1.29 0.2012
## Year1999 0.4140 0.1943 2.13 0.0373 *
## Year2000 0.3503 0.1898 1.85 0.0699 .
## Year2001 0.3728 0.1914 1.95 0.0561 .
## Year2002 0.2473 0.1471 1.68 0.0978 .
## Year2003 0.6041 0.3101 1.95 0.0561 .
## Year2004 0.2174 0.1914 1.14 0.2604
## Year2005 0.0963 0.1471 0.65 0.5151
## Year2006 0.3302 0.1854 1.78 0.0800 .
```

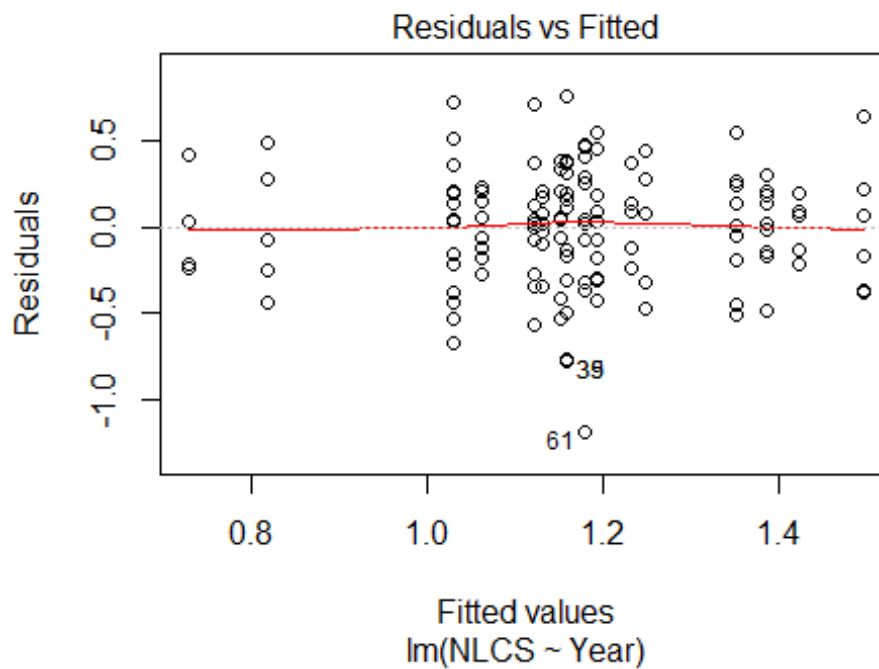


```

## Year2007          0.2255      0.6484      0.35      0.7292
## Year2008          0.5196      0.1831      2.84      0.0062 **
## Year2009          0.2471      0.2017      1.23      0.2253
## Year2010          0.7110      0.1557      4.57      2.5e-05 ***
## Year2011          0.2486      0.2046      1.21      0.2292
## Year2012          0.1259      0.2140      0.59      0.5587
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared:  0.231, Adjusted R-squared:  0.0131
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.313  0.894  0.970  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      1.28e-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: 78"
## [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
##   11   7  18  14  15  14   8   7  10   7   8  11   4  17  11
## 2011 2012
##   15   8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   6  14  11   9   5   5   4   7   4   6   8   4  15  10
## 2011 2012

```

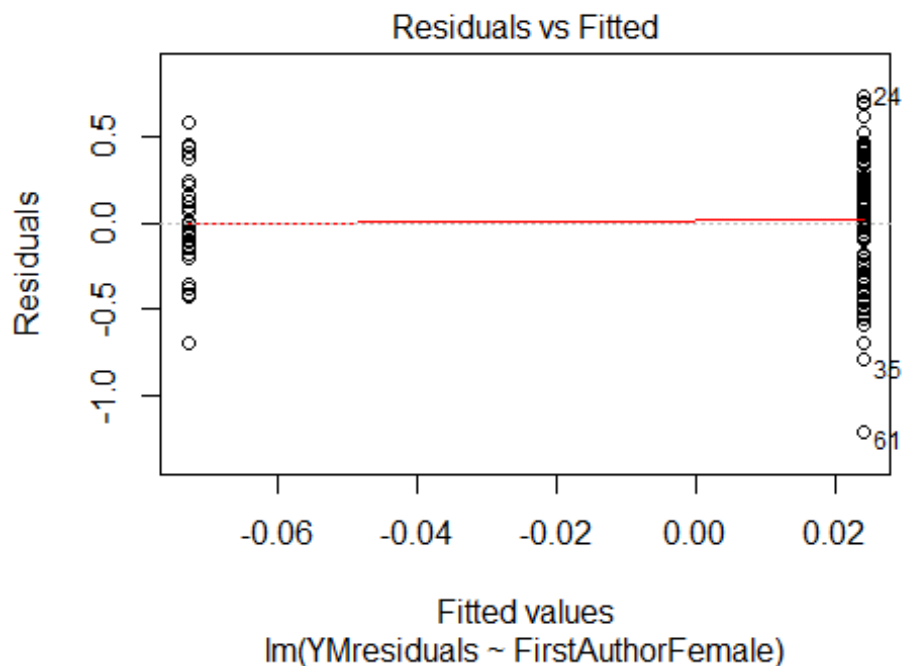
```
## 10 5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 6 14 7 7 5 5 4 6 4 6 8 4 15 9
## 2011 2012
## 9 5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## 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 = 1.4, df = 1, p-value = 0.2
## [1] "Female first author team size 2018 geometric mean: 4.89897948556636"
## [1] "Male first author team size 2018 geometric mean: 3.46694222261971"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

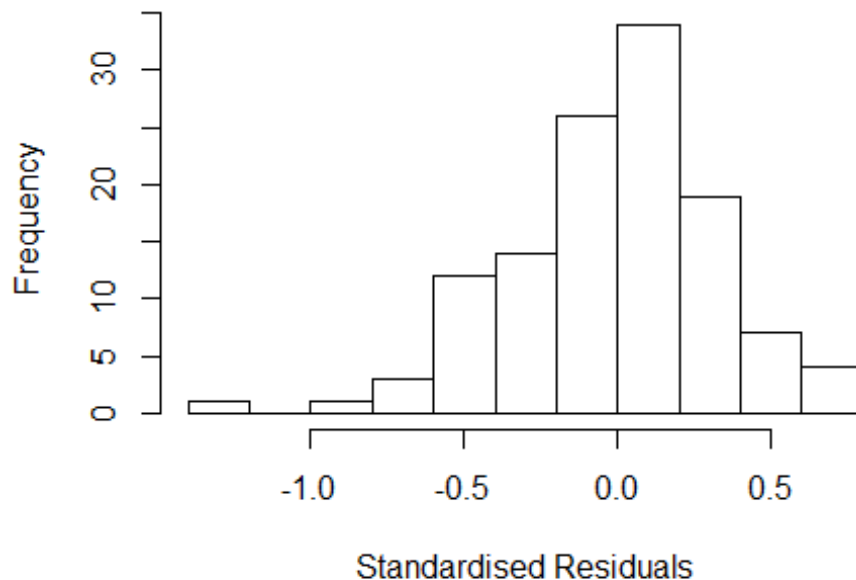
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.92820323027551"
## [1] "Male last author team size 2018 geometric mean: 3.23477147345614"

## 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 = 0.1
## 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  1.927  1      1.388
## LastAuthorFemale  1.790  1      1.338
## UniqueAuthors    14.650  4      1.399
## Year              26.591 16      1.108
```

Residuals from first and last author and team size



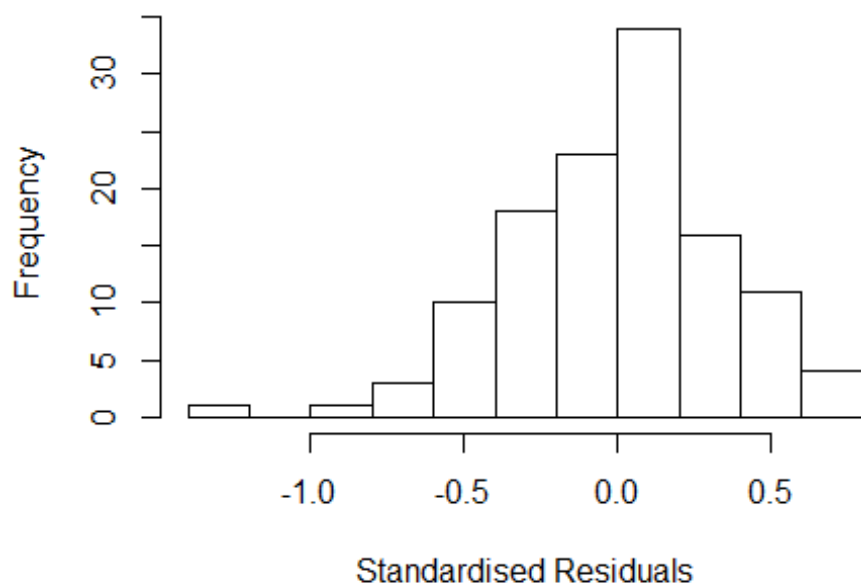
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2320 -0.2116 0.0181 0.1863 0.7169
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.46103 0.24902 5.87 6e-08 ***
## FirstAuthorFemale1 -0.10180 0.07898 -1.29 0.20050
## LastAuthorFemale1 -0.08371 0.06983 -1.20 0.23353
## UniqueAuthors2 -0.08639 0.21534 -0.40 0.68915
## UniqueAuthors3 -0.00658 0.23799 -0.03 0.97799
## UniqueAuthors4 -0.04062 0.26393 -0.15 0.87799
## UniqueAuthors5 -0.08772 0.23543 -0.37 0.71028
## Year1997 0.07222 0.19772 0.37 0.71571
## Year1998 -0.18250 0.19007 -0.96 0.33932
## Year1999 -0.18836 0.27956 -0.67 0.50202
```

```

## Year2000      0.00164    0.15399    0.01  0.99153
## Year2001     -0.52310    0.18353   -2.85  0.00533 **
## Year2002     -0.13713    0.23131   -0.59  0.55466
## Year2003     -0.21615    0.20917   -1.03  0.30397
## Year2004     -0.21588    0.14071   -1.53  0.12819
## Year2005     -0.17983    0.20503   -0.88  0.38258
## Year2006     -0.12758    0.16408   -0.78  0.43872
## Year2007     -0.27730    0.13926   -1.99  0.04924 *
## Year2008     -0.67175    0.17962   -3.74  0.00031 ***
## Year2009     -0.32666    0.16431   -1.99  0.04959 *
## Year2010     -0.17300    0.16400   -1.05  0.29408
## Year2011     -0.26493    0.16445   -1.61  0.11041
## Year2012      0.10361    0.11947    0.87  0.38791
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.347
## Multiple R-squared:  0.232, Adjusted R-squared:  0.06
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 111 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.182  0.884  0.954  0.916  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      8.26e-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 1.460 1      1.208
## LastAuthorFemale  1.466 1      1.211
## Year              2.088 16      1.023

```

Residuals from first and last author



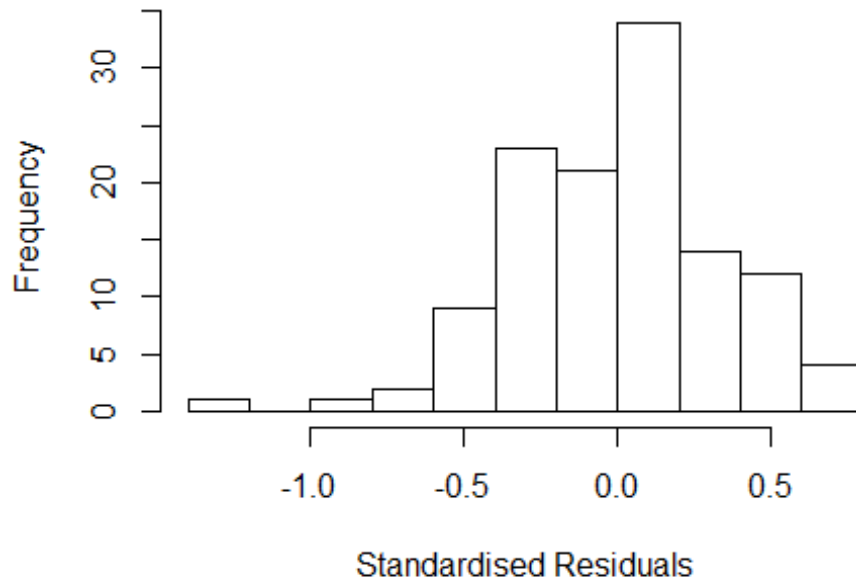
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2256 -0.2465  0.0355  0.2104  0.6785
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4129     0.1138   12.41 < 2e-16 ***
## FirstAuthorFemale1 -0.1100     0.0714   -1.54  0.12686
## LastAuthorFemale1 -0.0966     0.0669   -1.44  0.15210
## Year1997          0.0838     0.1937    0.43  0.66616
## Year1998         -0.1824     0.1797   -1.01  0.31256
## Year1999         -0.1873     0.2478   -0.76  0.45149
## Year2000          0.0131     0.1550    0.08  0.93295
## Year2001         -0.4922     0.1733   -2.84  0.00544 **
## Year2002         -0.1602     0.2143   -0.75  0.45658
## Year2003         -0.1811     0.2203   -0.82  0.41305
## Year2004         -0.2212     0.1478   -1.50  0.13752
## Year2005         -0.2022     0.2135   -0.95  0.34574
```

```

## Year2006          -0.1106      0.1611   -0.69   0.49406
## Year2007          -0.2476      0.1324   -1.87   0.06436 .
## Year2008          -0.6682      0.1722   -3.88   0.00018 ***
## Year2009          -0.3065      0.1660   -1.85   0.06776 .
## Year2010          -0.1547      0.1672   -0.92   0.35717
## Year2011          -0.2636      0.1621   -1.63   0.10697
## Year2012           0.0917      0.1269    0.72   0.47169
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.222, Adjusted R-squared:  0.0845
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 113 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.233 0.891 0.961 0.923 0.992 0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      8.26e-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 1.434 1      1.198
## Year              1.434 16      1.011

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2025 -0.2370 0.0194 0.1842 0.7074
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4000 0.1053 13.30 < 2e-16 ***
## FirstAuthorFemale1 -0.1198 0.0705 -1.70 0.09202 .
## Year1997 0.0987 0.1893 0.52 0.60330
## Year1998 -0.1738 0.1709 -1.02 0.31169
## Year1999 -0.1975 0.2356 -0.84 0.40382
## Year2000 0.0180 0.1458 0.12 0.90212
## Year2001 -0.5111 0.1812 -2.82 0.00575 **
## Year2002 -0.1472 0.2099 -0.70 0.48463
## Year2003 -0.2154 0.1960 -1.10 0.27426
## Year2004 -0.2198 0.1361 -1.62 0.10936
## Year2005 -0.2322 0.2179 -1.07 0.28915
## Year2006 -0.1112 0.1554 -0.72 0.47579
```

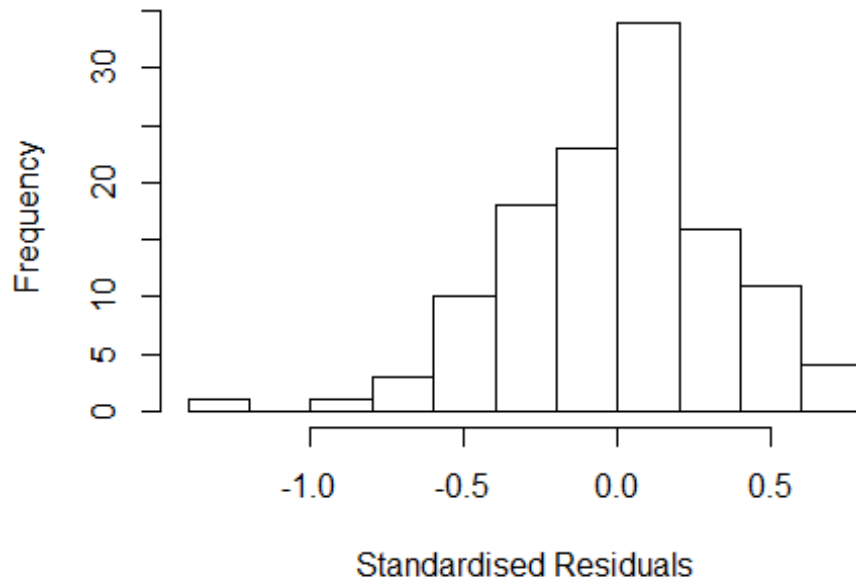


```

## Year2007          -0.2776      0.1200    -2.31  0.02269 *
## Year2008          -0.6816      0.1741    -3.91  0.00016 ***
## Year2009          -0.3305      0.1579    -2.09  0.03878 *
## Year2010          -0.1973      0.1587    -1.24  0.21663
## Year2011          -0.2834      0.1630    -1.74  0.08506 .
## Year2012           0.0699      0.1180     0.59  0.55481
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.362
## Multiple R-squared:  0.211, Adjusted R-squared:  0.0809
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 105 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.248  0.882  0.959  0.916  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      8.26e-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 1.433 1          1.197
## Year              1.433 16          1.011

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2273 -0.2324 0.0119 0.2398 0.7136
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4029 0.1249 11.23 < 2e-16 ***
## LastAuthorFemale1 -0.1074 0.0682 -1.57 0.11873
## Year1997 0.0703 0.2029 0.35 0.72961
## Year1998 -0.2075 0.1889 -1.10 0.27445
## Year1999 -0.1756 0.2452 -0.72 0.47559
## Year2000 -0.0517 0.1553 -0.33 0.73992
## Year2001 -0.5474 0.1918 -2.85 0.00523 **
## Year2002 -0.1503 0.2198 -0.68 0.49578
## Year2003 -0.1883 0.2101 -0.90 0.37208
## Year2004 -0.2462 0.1508 -1.63 0.10562
## Year2005 -0.1875 0.2180 -0.86 0.39178
## Year2006 -0.1538 0.1535 -1.00 0.31874
```

```

## Year2007          -0.2875      0.1442    -1.99   0.04876 *
## Year2008          -0.6553      0.1794    -3.65   0.00041 ***
## Year2009          -0.3267      0.1666    -1.96   0.05268 .
## Year2010          -0.1670      0.1761    -0.95   0.34530
## Year2011          -0.2499      0.1699    -1.47   0.14451
## Year2012           0.0623      0.1424     0.44   0.66272
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.209, Adjusted R-squared:  0.0781
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.234  0.906  0.958  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      8.26e-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: 121"
## [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]"
##
## 1998 2001 2002 2003 2004 2006 2007 2008
##    2    3    1    4    1    2    2    2
##
## 1998 2001 2002 2003 2004 2006 2007 2008
##    1    1    1    4    1    2    2    2
##
## 1998 2001 2002 2003 2004 2006 2007 2008
##    1    0    1    3    1    2    2    2

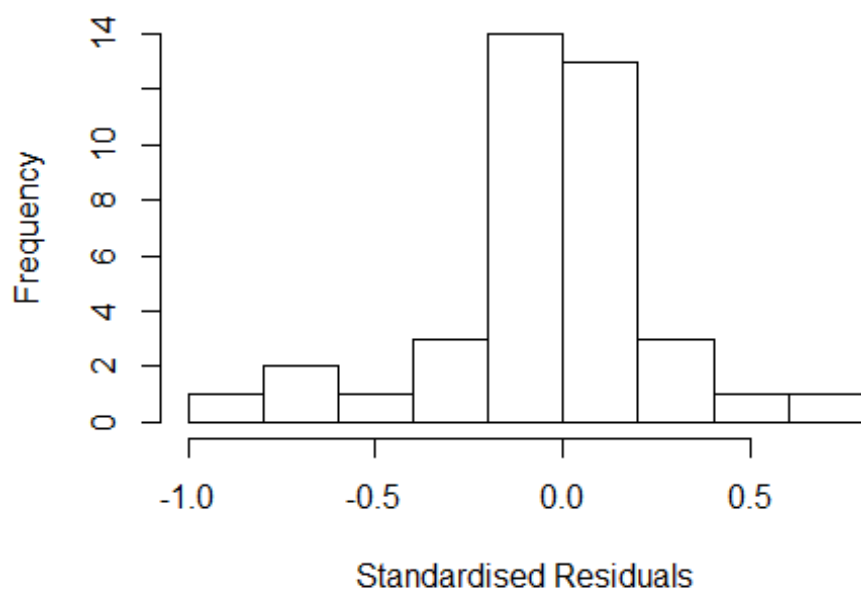
```

```

## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 12"
## [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
##      2      4      4      5      1      2      6      4      5      5      5      7      3      12      7
## 2011 2012
##      2      6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      3      1      4      1      1      5      2      3      2      4      4      1      8      4
## 2011 2012
##      2      3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      3      1      2      0      1      4      1      2      2      4      4      0      8      3
## 2011 2012
##      1      3
## [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: 5.65685424949238"
## [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"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.841e+12  1      1.960e+06
## LastAuthorFemale  2.105e+14  1      1.451e+07
## Year              1.075e+16 13      4.136e+00

```

Residuals from first and last author



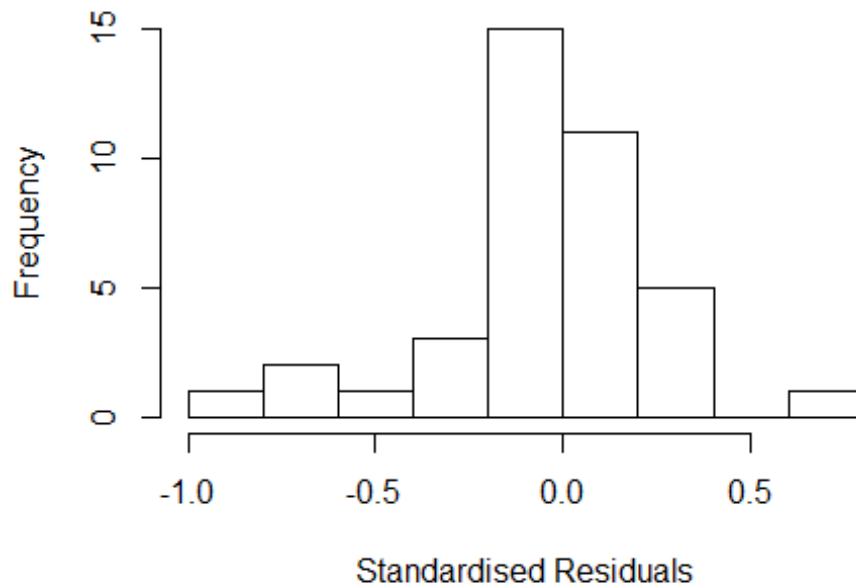
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.97e-01 -1.07e-01 1.39e-16 1.12e-01 7.12e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.11732 0.05232 21.36 < 2e-16 ***
## FirstAuthorFemale1 0.28597 0.13132 2.18 0.0399 *
## LastAuthorFemale1 0.02300 0.19591 0.12 0.9076
## Year1998 0.32268 0.05232 6.17 2.7e-06 ***
## Year1999 -0.16182 0.08499 -1.90 0.0695 .
## Year2001 -0.18832 0.05232 -3.60 0.0015 **
## Year2002 0.16198 0.23076 0.70 0.4898
## Year2003 -0.38432 0.05232 -7.35 1.8e-07 ***
## Year2004 -0.38180 0.04001 -9.54 1.8e-09 ***
## Year2005 -0.00228 0.12993 -0.02 0.9861
## Year2006 -0.20838 0.13400 -1.56 0.1336
## Year2007 0.03171 0.24492 0.13 0.8981
```

```

## Year2009          -0.08012    0.23638   -0.34    0.7377
## Year2010          -0.30326    0.14893   -2.04    0.0534 .
## Year2011           0.70668    0.05232   13.51   2.0e-12 ***
## Year2012          -0.06419    0.38580   -0.17    0.8693
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.258
## Multiple R-squared:  0.471, Adjusted R-squared:  0.126
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 31 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.202  0.873  0.968  0.875  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      2.56e-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 1.845e+15 1      4.296e+07
## Year              1.845e+15 13      3.865e+00

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -9.47e-01 -1.08e-01 -7.77e-16 9.05e-02 7.22e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.118219 0.060717 18.42 1.2e-15 ***
## FirstAuthorFemale1 0.283217 0.160578 1.76 0.0905 .
## Year1998 0.321781 0.060717 5.30 1.9e-05 ***
## Year1999 -0.162719 0.090759 -1.79 0.0856 .
## Year2001 -0.189219 0.060717 -3.12 0.0047 **
## Year2002 0.210380 0.114182 1.84 0.0778 .
## Year2003 -0.385219 0.060717 -6.34 1.5e-06 ***
## Year2004 -0.381328 0.043899 -8.69 7.1e-09 ***
## Year2005 -0.000436 0.144975 0.00 0.9976
## Year2006 -0.196674 0.108356 -1.82 0.0820 .
## Year2007 0.065718 0.214724 0.31 0.7622
## Year2009 -0.091364 0.284133 -0.32 0.7506
```

```

## Year2010          -0.306441    0.156746   -1.96    0.0623 .
## Year2011          0.705781    0.060717   11.62    2.4e-11 ***
## Year2012          0.017408    0.633065    0.03    0.9783
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.227
## Multiple R-squared:  0.54,   Adjusted R-squared:  0.272
## Convergence in 34 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 31 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0442 0.8530 0.9660 0.8440 0.9870 0.9930
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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 2.214 1          1.488
## Year              2.214 13          1.031
##
## [1] "List of 0 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
## -0.864 -0.118  0.037  0.149  0.558
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)

```



```

## (Intercept)          1.2098      0.1103    10.97  7.9e-11 ***
## LastAuthorFemale1    0.0168      0.2115     0.08  0.93743
## Year1998              0.2302      0.1103     2.09  0.04776 *
## Year1999             -0.2543      0.1286    -1.98  0.05963 .
## Year2001             -0.2808      0.1103    -2.55  0.01776 *
## Year2002              0.0357      0.3084     0.12  0.90881
## Year2003             -0.4768      0.1103    -4.32  0.00023 ***
## Year2004             -0.3313      0.1583    -2.09  0.04715 *
## Year2005              0.1912      0.1407     1.36  0.18675
## Year2006             -0.3018      0.1630    -1.85  0.07638 .
## Year2007             -0.0975      0.2200    -0.44  0.66157
## Year2009             -0.0189      0.1822    -0.10  0.91840
## Year2010             -0.2870      0.1705    -1.68  0.10521
## Year2011              0.6142      0.1103     5.57  1.0e-05 ***
## Year2012             -0.2054      0.3446    -0.60  0.55673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.312, Adjusted R-squared:  -0.0891
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 32 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.570  0.909  0.984  0.931  0.994  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           2.56e-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: 39"
## [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]"

```

```

##
## 1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    2    2    1    1    1    2    7    2    1    1    1    3    2
##
## 1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    0    0    1    0    1    1    1    2    1    1    1    0    2    2
##
## 1997 1998 1999 2000 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    0    0    0    0    0    1    1    2    1    1    1    0    2    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: 5"
## [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: 11"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    3    9    3    2    2    1    5    5    1    5    3    4    6    4    4
## 2012
##    6
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    0    1    1    0    1    2    0    1    2    0    2    1    3
## 2012
##    3
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    0    1    1    0    1    2    0    1    2    0    1    1    2
## 2012
##    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 2.91295063024394"

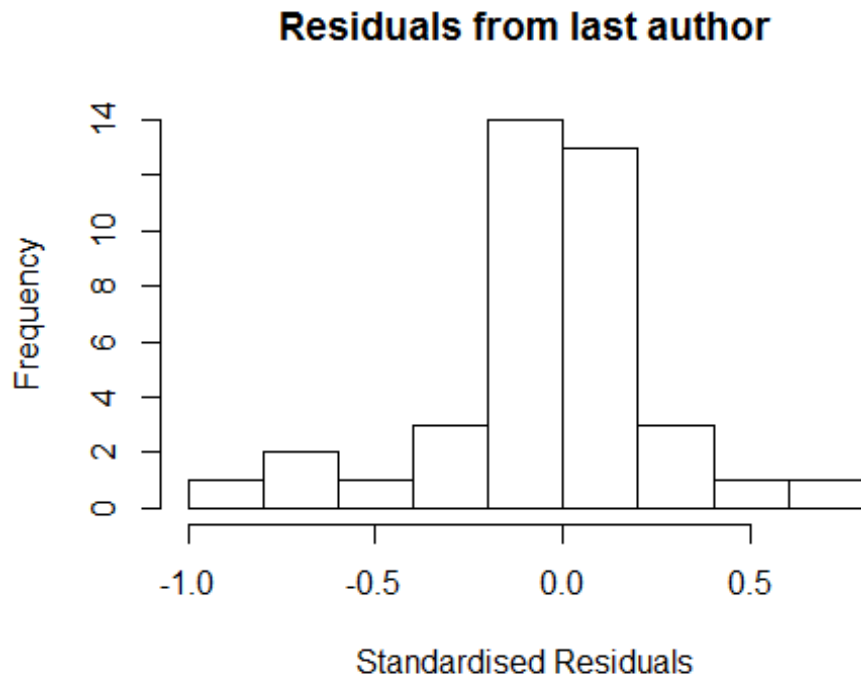
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```

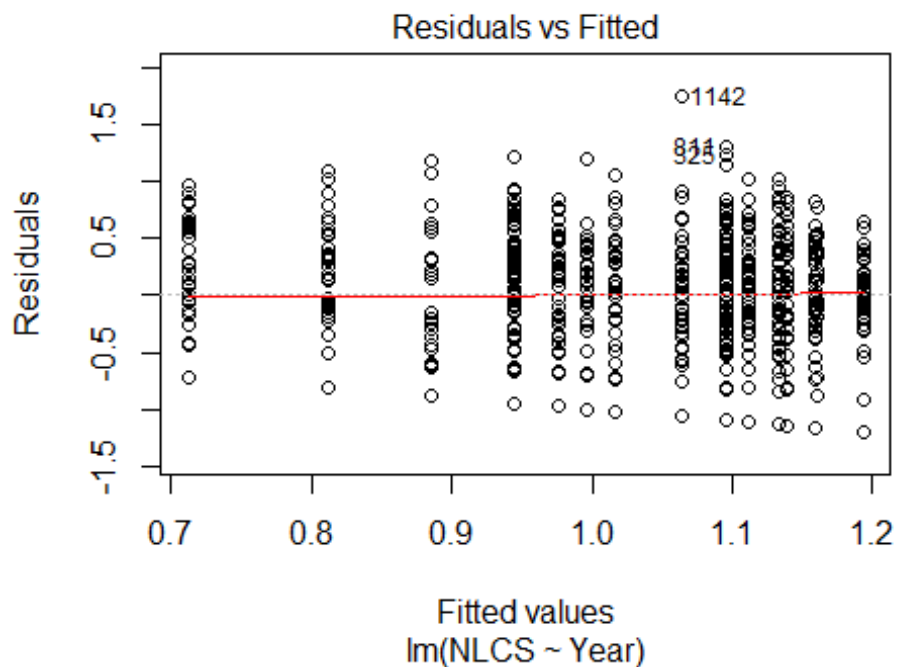
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 2.93015605158352"
## [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: 17"
## [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
##      2      3      5      1      2      4      3      1      3      1      1      4      3      4      4
## 2011 2012
##      4      4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      0      2      1      1      3      0      1      2      0      0      3      2      2      4
## 2011 2012
##      1      1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      0      2      1      0      3      0      1      2      0      0      2      2      2      4
## 2011 2012
##      1      1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 3.22370979547063"
## 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 = 2.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 3.17767152314644"
## [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: 23"
## [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
## 58 77 80 83 71 56 58 41 57 50 64 76 82 88 72
## 2011 2012
```

```
## 73 62
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 40 57 59 59 41 20 39 28 38 32 44 48 48 51 45
## 2011 2012
## 48 34
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 39 53 53 52 36 17 34 25 33 29 35 43 37 43 38
## 2011 2012
## 45 29
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 12, df = 16, p-value = 0.7
```



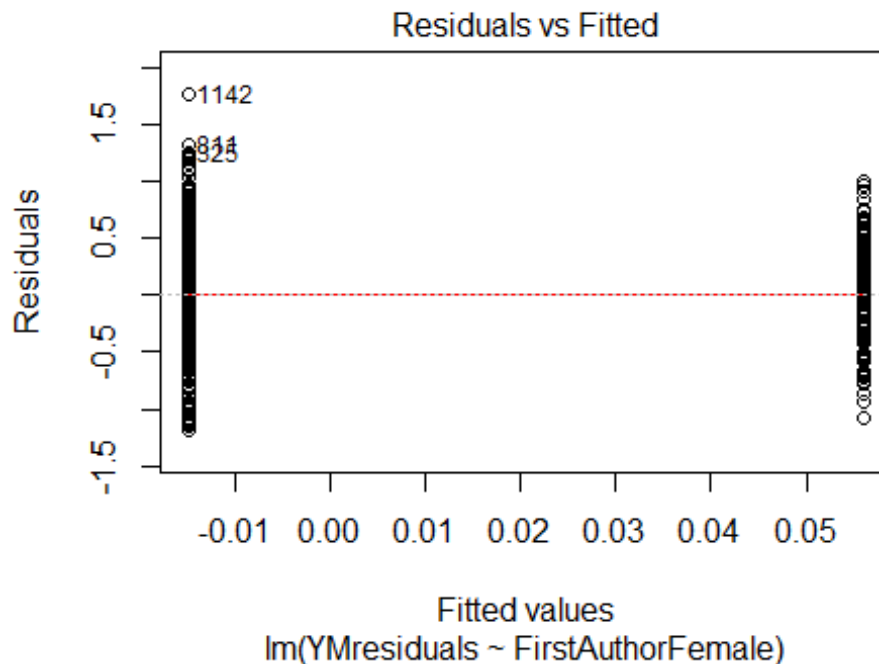
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.5, df = 1, p-value = 0.06

## [1] "Female first author team size 2018 geometric mean: 3.85548753045898"
## [1] "Male first author team size 2018 geometric mean: 3.23601081369347"
```

```
## 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 = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.75235359512996"
## [1] "Male last author team size 2018 geometric mean: 3.17677644561033"

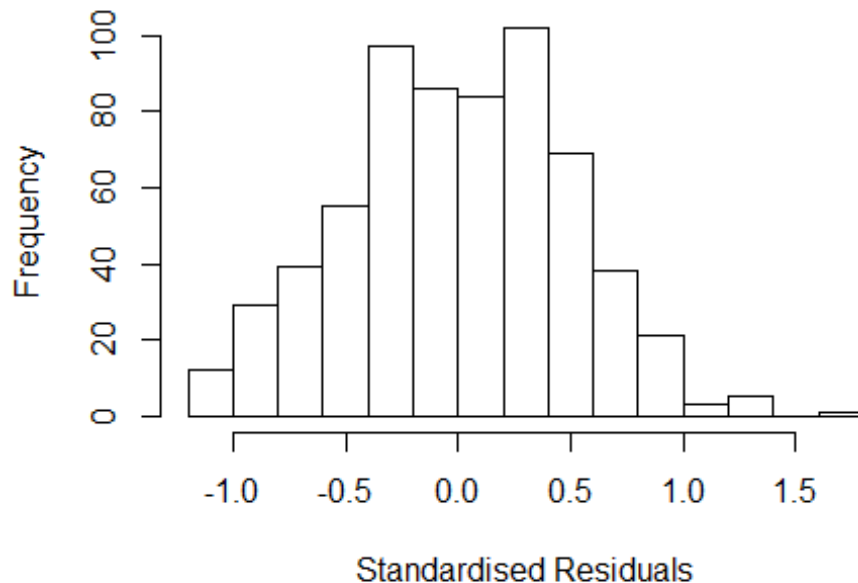
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120, 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 1.187 1          1.090
```

## LastAuthorFemale	1.158	1	1.076
## UniqueAuthors	1.617	4	1.062
## Year	1.792	16	1.018

Residuals from first and last author and team size



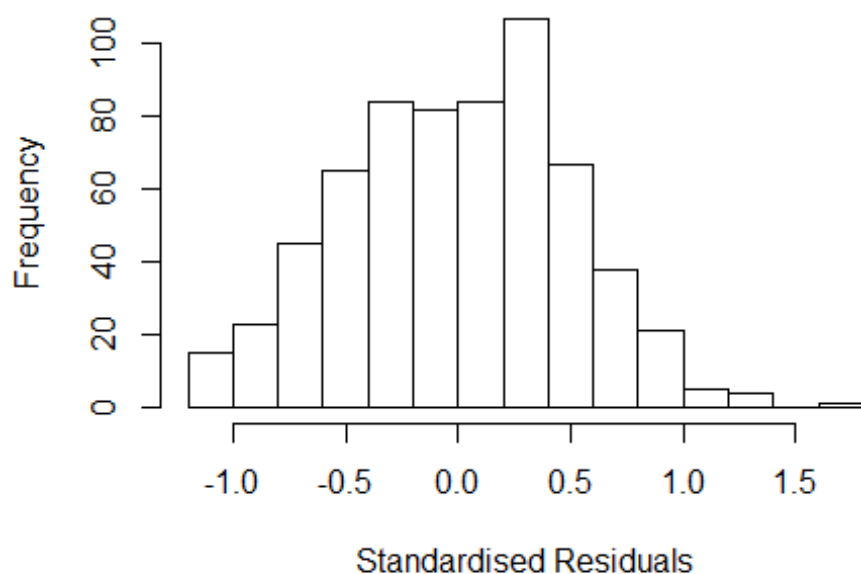
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11658 -0.34424 0.00701 0.35499 1.77989
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0842 0.1118 9.70 <2e-16 ***
## FirstAuthorFemale1 0.0539 0.0494 1.09 0.2757
## LastAuthorFemale1 0.0210 0.0515 0.41 0.6843
## UniqueAuthors2 0.0405 0.0777 0.52 0.6026
## UniqueAuthors3 -0.0509 0.0795 -0.64 0.5228
## UniqueAuthors4 0.0545 0.0876 0.62 0.5342
```

```

## UniqueAuthors5      0.1995      0.0861      2.32      0.0208 *
## Year1997             -0.0368      0.1130     -0.33      0.7450
## Year1998             -0.0028      0.1169     -0.02      0.9809
## Year1999             -0.0554      0.1273     -0.43      0.6638
## Year2000              0.0229      0.1111      0.21      0.8370
## Year2001              0.0695      0.1587      0.44      0.6618
## Year2002              0.0324      0.1178      0.27      0.7836
## Year2003             -0.2575      0.1575     -1.63      0.1027
## Year2004             -0.1587      0.1359     -1.17      0.2436
## Year2005             -0.0330      0.1360     -0.24      0.8081
## Year2006             -0.4278      0.1405     -3.05      0.0024 **
## Year2007             -0.2216      0.1257     -1.76      0.0783 .
## Year2008             -0.3946      0.1247     -3.16      0.0016 **
## Year2009             -0.1096      0.1163     -0.94      0.3465
## Year2010             -0.2715      0.1256     -2.16      0.0310 *
## Year2011             -0.1894      0.1238     -1.53      0.1266
## Year2012             -0.1993      0.1354     -1.47      0.1418
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.512
## Multiple R-squared:  0.0904, Adjusted R-squared:  0.058
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 45 weights are ~= 1. The remaining 596 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.201  0.881  0.952  0.916  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      1.56e-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 1.128 1      1.062
## LastAuthorFemale  1.137 1      1.066
## Year              1.218 16      1.006

```


Residuals from first and last author



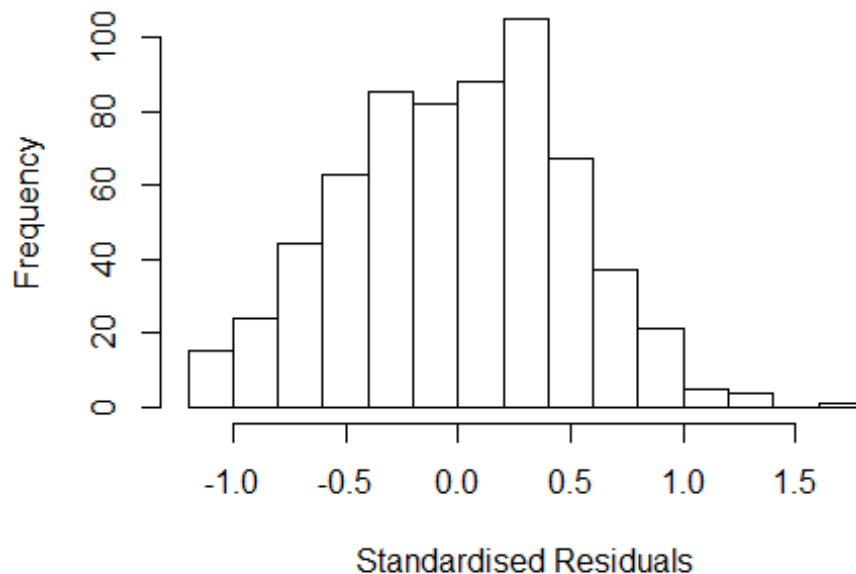
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1389 -0.3786 0.0148 0.3548 1.7909
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.138944 0.093475 12.18 <2e-16 ***
## FirstAuthorFemale1 0.070413 0.048546 1.45 0.1474
## LastAuthorFemale1 0.025811 0.050792 0.51 0.6115
## Year1997 -0.065264 0.112074 -0.58 0.5606
## Year1998 -0.028372 0.114302 -0.25 0.8040
## Year1999 -0.077123 0.125657 -0.61 0.5396
## Year2000 -0.002328 0.109223 -0.02 0.9830
## Year2001 -0.008521 0.157546 -0.05 0.9569
## Year2002 -0.000153 0.118235 0.00 0.9990
## Year2003 -0.296670 0.150165 -1.98 0.0486 *
## Year2004 -0.194089 0.136247 -1.42 0.1548
## Year2005 -0.077087 0.137178 -0.56 0.5744
```

```

## Year2006          -0.448420    0.138644   -3.23    0.0013 **
## Year2007          -0.230750    0.129533   -1.78    0.0753 .
## Year2008          -0.397065    0.124891   -3.18    0.0015 **
## Year2009          -0.120860    0.118218   -1.02    0.3070
## Year2010          -0.280355    0.129368   -2.17    0.0306 *
## Year2011          -0.206165    0.122495   -1.68    0.0929 .
## Year2012          -0.186878    0.136338   -1.37    0.1710
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.523
## Multiple R-squared:  0.0676, Adjusted R-squared:  0.0406
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 49 weights are ~= 1. The remaining 592 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.218  0.882  0.951  0.917  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      1.56e-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 1.106 1          1.052
## Year              1.106 16          1.003

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.1434 -0.3808 0.0135 0.3508 1.7867
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.14151 0.09389 12.16 <2e-16 ***
## FirstAuthorFemale1 0.07274 0.04823 1.51 0.1320
## Year1997 -0.06668 0.11250 -0.59 0.5536
## Year1998 -0.02991 0.11485 -0.26 0.7946
## Year1999 -0.07724 0.12602 -0.61 0.5402
## Year2000 0.00190 0.10925 0.02 0.9862
## Year2001 -0.00760 0.15841 -0.05 0.9617
## Year2002 0.00131 0.11862 0.01 0.9912
## Year2003 -0.29466 0.15087 -1.95 0.0513 .
## Year2004 -0.19374 0.13743 -1.41 0.1591
## Year2005 -0.07189 0.13671 -0.53 0.5992
## Year2006 -0.44881 0.13909 -3.23 0.0013 **
```

```

## Year2007          -0.22830    0.12959   -1.76    0.0786 .
## Year2008          -0.39476    0.12459   -3.17    0.0016 **
## Year2009          -0.11919    0.11849   -1.01    0.3148
## Year2010          -0.27667    0.12924   -2.14    0.0327 *
## Year2011          -0.20476    0.12284   -1.67    0.0961 .
## Year2012          -0.18489    0.13668   -1.35    0.1766
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.523
## Multiple R-squared:  0.0673, Adjusted R-squared:  0.0419
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 52 weights are ~= 1. The remaining 589 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.218  0.880  0.951  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      1.56e-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 1.107 1          1.052
## 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.1523 -0.3825  0.0197  0.3622  1.7732

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.14694    0.09323   12.30  <2e-16 ***
## LastAuthorFemale1  0.03342    0.05082    0.66  0.5110
## Year1997       -0.06390    0.11175   -0.57  0.5677
## Year1998       -0.01933    0.11341   -0.17  0.8647
## Year1999       -0.06614    0.12473   -0.53  0.5961
## Year2000        0.00535    0.10949    0.05  0.9610
## Year2001       -0.00508    0.15508   -0.03  0.9739
## Year2002        0.00401    0.11829    0.03  0.9730
## Year2003       -0.29608    0.15013   -1.97  0.0490 *
## Year2004       -0.18247    0.13658   -1.34  0.1820
## Year2005       -0.07866    0.13817   -0.57  0.5693
## Year2006       -0.44102    0.14062   -3.14  0.0018 **
## Year2007       -0.23308    0.12915   -1.80  0.0716 .
## Year2008       -0.38992    0.12507   -3.12  0.0019 **
## Year2009       -0.11115    0.11755   -0.95  0.3447
## Year2010       -0.27109    0.12817   -2.12  0.0348 *
## Year2011       -0.19705    0.12200   -1.62  0.1068
## Year2012       -0.17452    0.13654   -1.28  0.2017
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.532
## Multiple R-squared:  0.0643, Adjusted R-squared:  0.0388
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 45 weights are ~= 1. The remaining 596 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.245  0.885   0.951   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           1.56e-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: 641"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
##    4    4    1    2    9    4    3    2    5    3    9    3    1    2    5
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
##    2    2    1    2    3    3    1    0    2    3    2    2    0    2    2
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011 2012
##    2    2    1    1    3    1    0    0    1    2    2    2    0    1    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.23606797749979"
## [1] "Male first author team size 2018 geometric mean: 4"

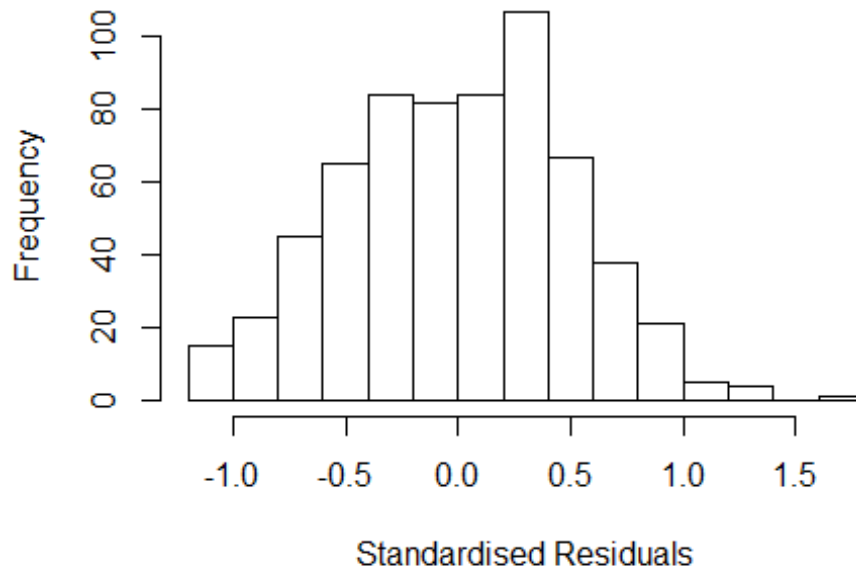
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 4.64158883361278"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

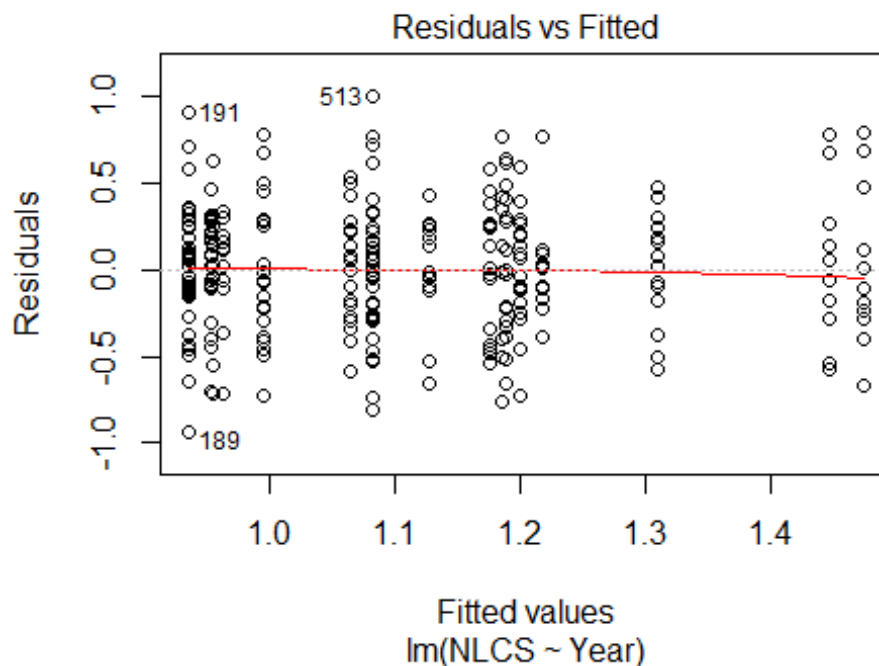
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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,
## [1] "Regression 2: First author gender, Last author gender, Year as
## [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: 20"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1602"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
## [1] "gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 24 22 17 23 19 15 20 20 21 23 34 32 22 35 28
## 2011 2012
## 35 48
##
```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 15 8 12 11 9 13 15 11 18 23 19 15 21 17
## 2011 2012
## 21 34
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 13 7 9 9 7 13 12 10 18 23 18 13 18 13
## 2011 2012
## 19 29
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.6
```



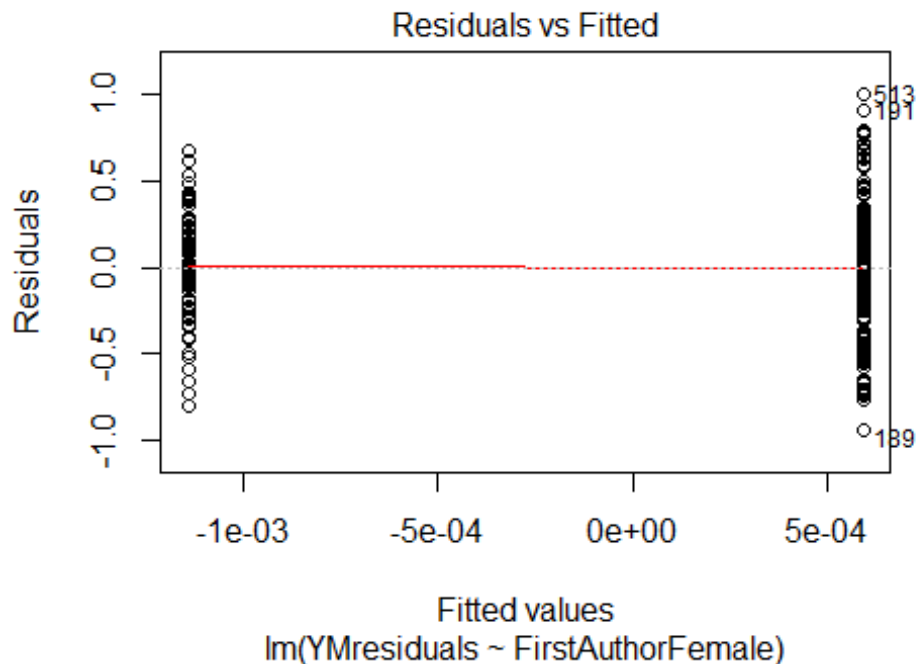
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.8, df = 1, p-value = 0.002
## [1] "Female first author team size 2018 geometric mean: 3.80675409583932"
## [1] "Male first author team size 2018 geometric mean: 2.74661447899667"
```



```
## 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 = 36, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.81712059283214"
## [1] "Male last author team size 2018 geometric mean: 3.34055403045353"

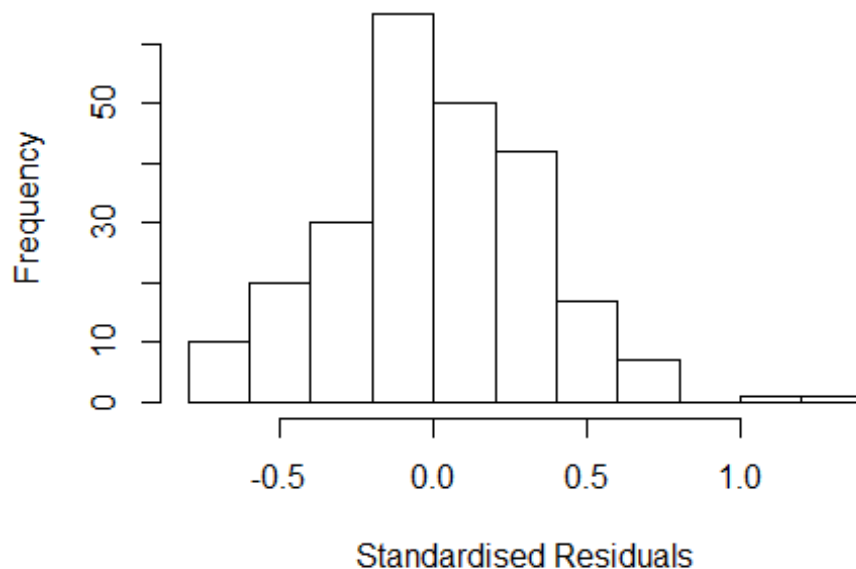
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7.5, 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 1.749 1          1.323
```

## LastAuthorFemale	1.890	1	1.375
## UniqueAuthors	3.322	4	1.162
## Year	7.071	16	1.063

Residuals from first and last author and team size



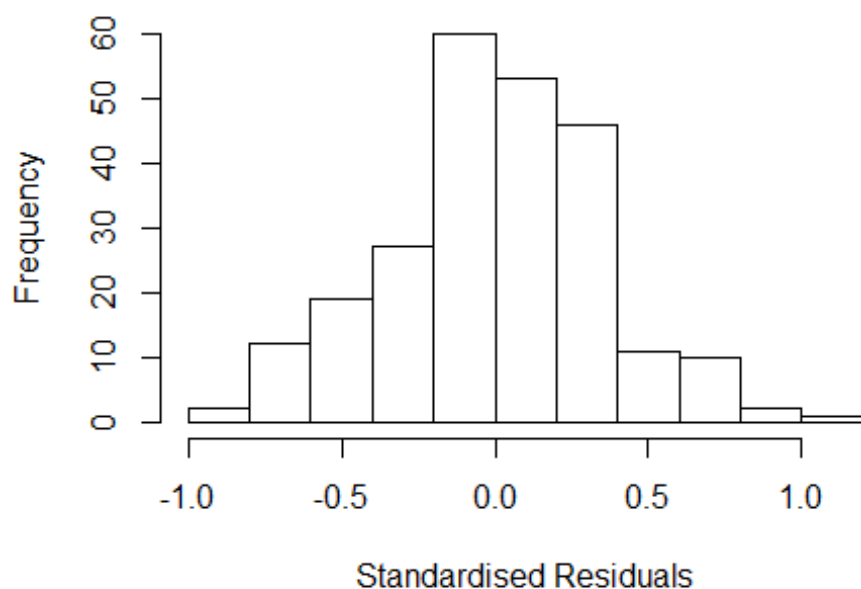
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.79106 -0.19780 -0.00941 0.22598 1.21666
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9890 0.1265 7.81 2.2e-13 ***
## FirstAuthorFemale1 0.0228 0.0496 0.46 0.64580
## LastAuthorFemale1 -0.0081 0.0716 -0.11 0.90999
## UniqueAuthors2 0.1598 0.1337 1.20 0.23329
## UniqueAuthors3 0.2586 0.1309 1.98 0.04938 *
## UniqueAuthors4 0.1542 0.1310 1.18 0.24029
```

```

## UniqueAuthors5      0.3030      0.1266      2.39  0.01755 *
## Year1997             0.1357      0.1027      1.32  0.18766
## Year1998             0.1089      0.2350      0.46  0.64355
## Year1999             0.1295      0.1274      1.02  0.31037
## Year2000             0.0254      0.1280      0.20  0.84319
## Year2001            -0.1822      0.1816     -1.00  0.31686
## Year2002            -0.0940      0.1137     -0.83  0.40917
## Year2003            -0.3740      0.1066     -3.51  0.00055 ***
## Year2004            -0.2224      0.1061     -2.10  0.03711 *
## Year2005             0.0110      0.0960      0.11  0.90872
## Year2006            -0.2188      0.0947     -2.31  0.02177 *
## Year2007            -0.1419      0.1060     -1.34  0.18180
## Year2008            -0.0733      0.1283     -0.57  0.56835
## Year2009            -0.2135      0.1159     -1.84  0.06676 .
## Year2010            -0.1217      0.1433     -0.85  0.39685
## Year2011            -0.2789      0.1004     -2.78  0.00594 **
## Year2012            -0.2034      0.1070     -1.90  0.05859 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.312
## Multiple R-squared:  0.18,   Adjusted R-squared:  0.0981
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 221 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0951 0.8540 0.9470 0.8930 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      4.12e-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 1.599 1      1.264
## LastAuthorFemale  1.612 1      1.270
## Year              2.337 16      1.027

```

Residuals from first and last author



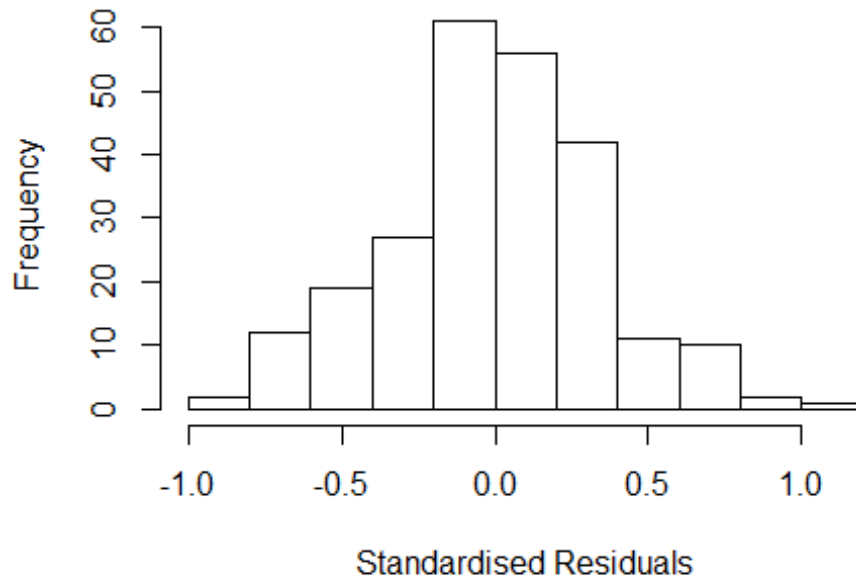
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.86179 -0.18438  0.00749  0.22442  1.08032
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1687    0.0645   18.11 < 2e-16 ***
## FirstAuthorFemale1  0.0332    0.0487    0.68  0.49641
## LastAuthorFemale1 -0.0204    0.0720   -0.28  0.77765
## Year1997          0.1467    0.1027    1.43  0.15452
## Year1998          0.1191    0.2542    0.47  0.63991
## Year1999          0.1331    0.1195    1.11  0.26684
## Year2000          0.1136    0.1321    0.86  0.39071
## Year2001         -0.2076    0.2257   -0.92  0.35887
## Year2002         -0.0351    0.1038   -0.34  0.73575
## Year2003         -0.3615    0.1048   -3.45  0.00067 ***
## Year2004         -0.2170    0.1096   -1.98  0.04897 *
## Year2005          0.0294    0.0951    0.31  0.75723
```

```

## Year2006          -0.2038      0.0924    -2.21  0.02838 *
## Year2007          -0.0940      0.1021    -0.92  0.35846
## Year2008          -0.0614      0.1333    -0.46  0.64564
## Year2009          -0.1736      0.1220    -1.42  0.15620
## Year2010          -0.0353      0.1228    -0.29  0.77417
## Year2011          -0.2430      0.0944    -2.57  0.01074 *
## Year2012          -0.1447      0.0984    -1.47  0.14283
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.138, Adjusted R-squared:  0.069
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 225 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
## 0.236 0.857 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      4.12e-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 1.477 1      1.215
## Year      1.477 16      1.012

```

Residuals from first author



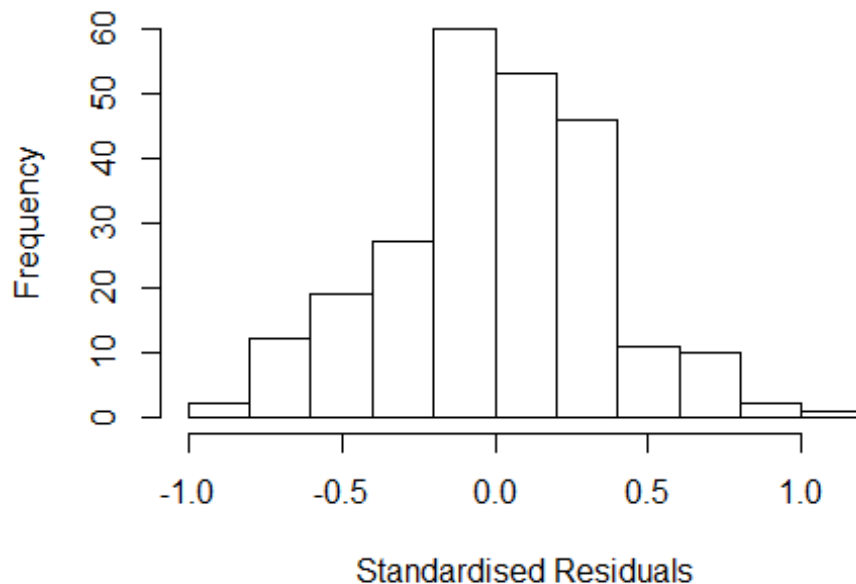
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.85937 -0.18884 0.00666 0.22502 1.06298
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1663 0.0630 18.52 < 2e-16 ***
## FirstAuthorFemale1 0.0320 0.0474 0.67 0.50077
## Year1997 0.1450 0.1027 1.41 0.15940
## Year1998 0.1190 0.2505 0.48 0.63511
## Year1999 0.1340 0.1165 1.15 0.25155
## Year2000 0.1136 0.1322 0.86 0.39095
## Year2001 -0.2067 0.2230 -0.93 0.35499
## Year2002 -0.0336 0.1019 -0.33 0.74216
## Year2003 -0.3634 0.1038 -3.50 0.00056 ***
## Year2004 -0.2152 0.1069 -2.01 0.04523 *
## Year2005 0.0294 0.0941 0.31 0.75518
## Year2006 -0.2026 0.0895 -2.26 0.02458 *
```

```

## Year2007          -0.0990      0.0992   -1.00  0.31971
## Year2008          -0.0649      0.1325   -0.49  0.62491
## Year2009          -0.1759      0.1213   -1.45  0.14843
## Year2010          -0.0422      0.1200   -0.35  0.72553
## Year2011          -0.2438      0.0943   -2.59  0.01033 *
## Year2012          -0.1453      0.0981   -1.48  0.14004
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.327
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0718
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 224 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.270  0.866  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      4.12e-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 1.559 1      1.249
## Year              1.559 16      1.014

```

Residuals from last author



```
## [1] "List of 0 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
## -0.87728 -0.19256 0.00613 0.22161 1.06269
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1818 0.0628 18.81 < 2e-16 ***
## LastAuthorFemale1 -0.0175 0.0705 -0.25 0.80435
## Year1997 0.1390 0.1021 1.36 0.17496
## Year1998 0.1215 0.2547 0.48 0.63380
## Year1999 0.1292 0.1207 1.07 0.28576
## Year2000 0.1064 0.1312 0.81 0.41826
## Year2001 -0.2217 0.2240 -0.99 0.32355
## Year2002 -0.0314 0.1044 -0.30 0.76398
## Year2003 -0.3619 0.1051 -3.44 0.00069 ***
## Year2004 -0.2228 0.1103 -2.02 0.04466 *
## Year2005 0.0306 0.0960 0.32 0.75053
## Year2006 -0.2019 0.0931 -2.17 0.03115 *
```

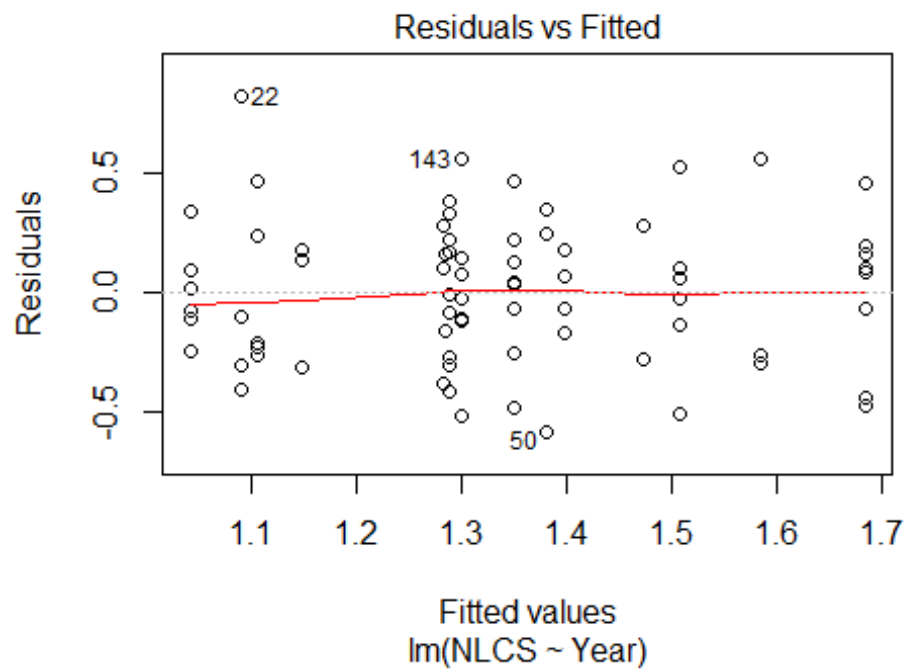


```

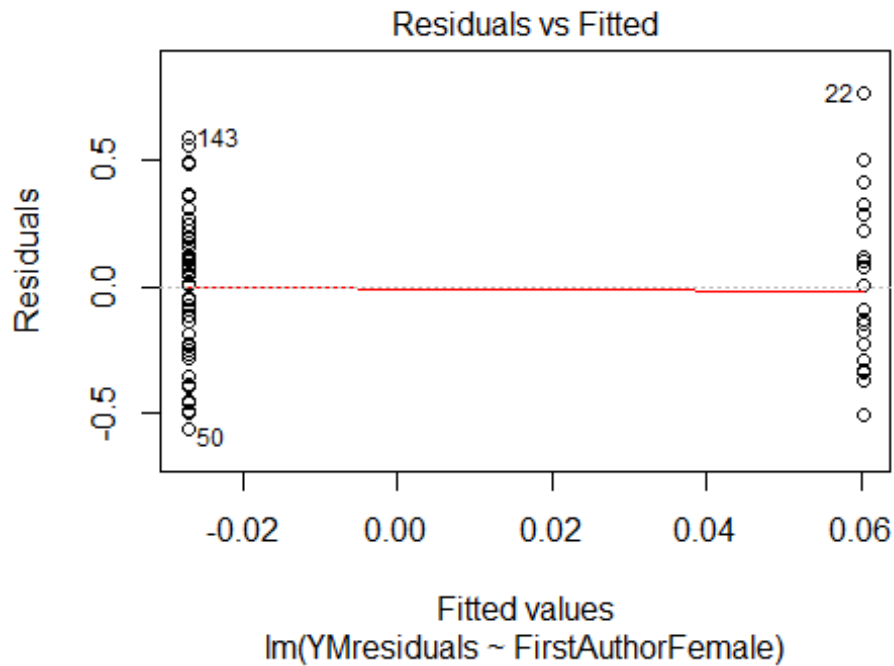
## Year2007          -0.0944      0.1013    -0.93   0.35241
## Year2008          -0.0581      0.1332    -0.44   0.66306
## Year2009          -0.1799      0.1189    -1.51   0.13150
## Year2010          -0.0395      0.1226    -0.32   0.74755
## Year2011          -0.2422      0.0950    -2.55   0.01148 *
## Year2012          -0.1430      0.0990    -1.44   0.14995
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.325
## Multiple R-squared:  0.136, Adjusted R-squared:  0.0711
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 225 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.264  0.865   0.955   0.898   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      4.12e-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: 243"
## [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
##    4   12    8    9    5    4    4    1    5    7    8   13    5    9    8
## 2011 2012
##    9    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    8    4    6    3    2    2    0    3    4    5    9    3    9    7
## 2011 2012

```

```
##      6      3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      6      2      4      3      2      1      0      2      3      3      5      0      8      2
## 2011 2012
##      2      2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 8.4, df = 14, p-value = 0.9
```

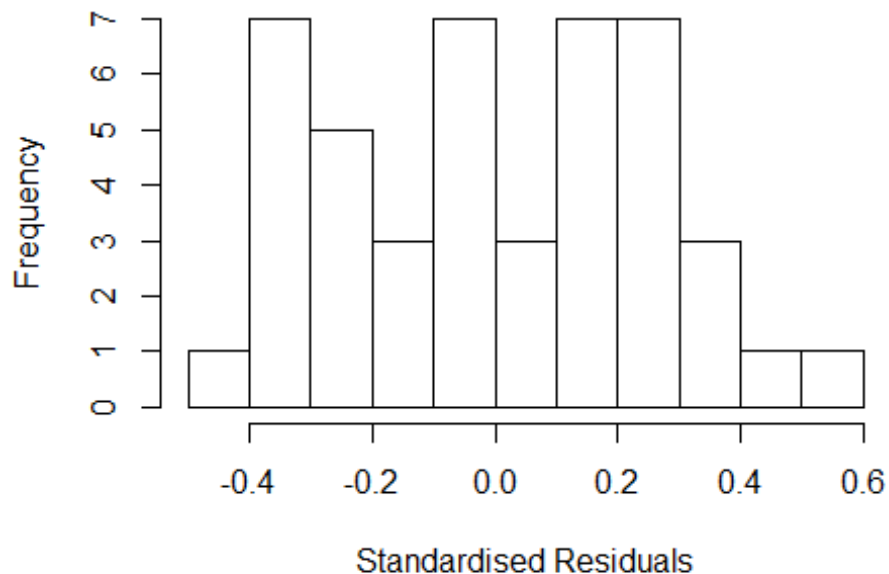


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.17, df = 1, p-value = 0.7
```



```
## [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"
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -5.786e+14  1          NaN
## LastAuthorFemale  7.430e+00  1          2.726
## UniqueAuthors    -1.922e+17  4          NaN
## Year              -5.149e+17 13          NaN
```

Residuals from first and last author and team size



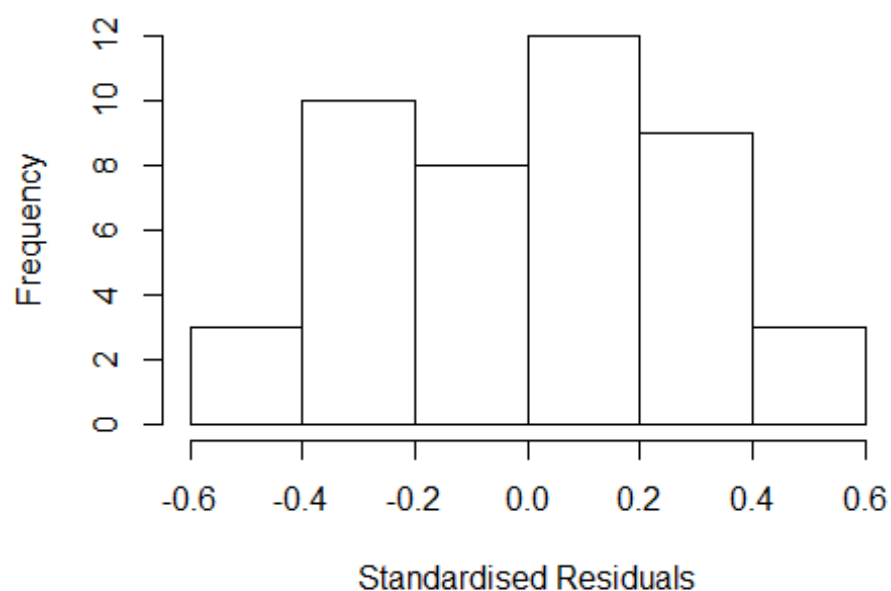
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -4.18e-01 -2.35e-01 2.22e-16 2.01e-01 5.43e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.98380 0.20969 9.46 9.7e-10 ***
## FirstAuthorFemale1 0.17359 0.10938 1.59 0.12507
## LastAuthorFemale1 -0.19052 0.11977 -1.59 0.12424
## UniqueAuthors2 -0.36536 0.17299 -2.11 0.04485 *
## UniqueAuthors3 -0.38314 0.15524 -2.47 0.02078 *
## UniqueAuthors4 -0.14932 0.18968 -0.79 0.43854
## UniqueAuthors5 -0.80361 0.22104 -3.64 0.00125 **
## Year1998 -0.78695 0.17713 -4.44 0.00016 ***
## Year1999 -0.23351 0.19456 -1.20 0.24131
## Year2000 -0.28130 0.26292 -1.07 0.29488
```

```

## Year2001          -0.13548      0.26187      -0.52      0.60945
## Year2002          -0.15266      0.16835      -0.91      0.37317
## Year2004           0.12252      0.38162       0.32      0.75084
## Year2005          -0.29426      0.21165      -1.39      0.17668
## Year2006          -1.05238      0.19159      -5.49      1.0e-05 ***
## Year2007          -0.19622      0.21425      -0.92      0.36849
## Year2009          -0.33978      0.16821      -2.02      0.05423 .
## Year2010           0.00421      0.23895       0.02      0.98610
## Year2011          -0.86976      0.21770      -4.00      0.00050 ***
## Year2012          -0.60788      0.19014      -3.20      0.00374 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.447
## Multiple R-squared:  0.49,   Adjusted R-squared:  0.102
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.870  0.946  0.975  0.966  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           2.22e-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 7.762 1 2.786
## LastAuthorFemale 2.461 1 1.569
## Year 11.234 13 1.098

```

Residuals from first and last author



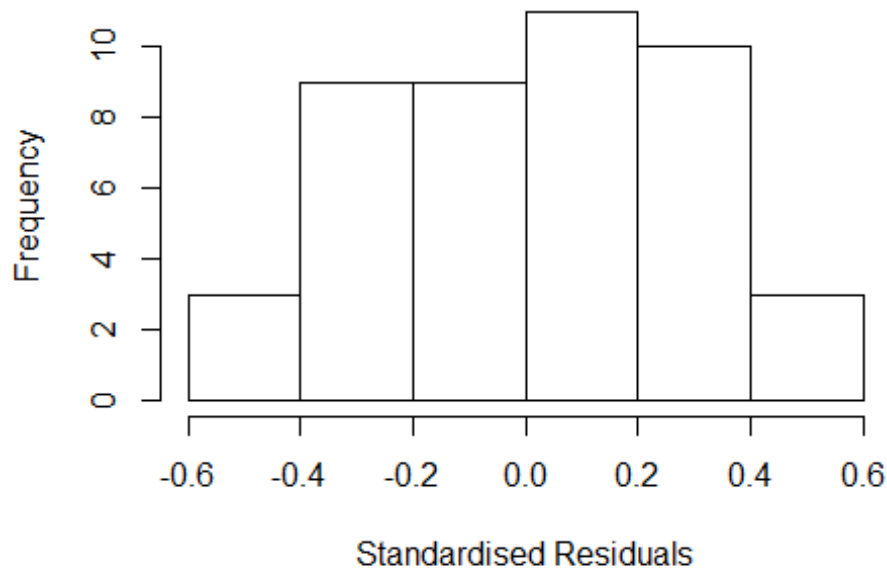
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.52549 -0.22555 0.00747 0.20039 0.53847
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.605 0.152 10.58 1.8e-11 ***
## FirstAuthorFemale1 0.145 0.116 1.25 0.22213
## LastAuthorFemale1 -0.100 0.145 -0.69 0.49623
## Year1998 -0.774 0.189 -4.10 0.00031 ***
## Year1999 -0.224 0.193 -1.16 0.25476
## Year2000 -0.183 0.273 -0.67 0.50850
## Year2001 -0.153 0.253 -0.61 0.54868
## Year2002 -0.158 0.152 -1.04 0.30786
## Year2004 0.104 0.372 0.28 0.78078
## Year2005 -0.294 0.190 -1.54 0.13324
## Year2006 -0.833 0.157 -5.30 1.1e-05 ***
## Year2007 -0.170 0.228 -0.75 0.46097
```

```

## Year2009          -0.376      0.186   -2.02  0.05266 .
## Year2010          -0.037      0.289   -0.13  0.89898
## Year2011          -0.735      0.158   -4.63  7.0e-05 ***
## Year2012          -0.621      0.215   -2.88  0.00732 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.392
## Multiple R-squared:  0.428, Adjusted R-squared:  0.133
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.836  0.925  0.970  0.953  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      2.22e-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 7.445 1          2.728
## Year              7.445 13          1.080

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.5763 -0.2234 0.0255 0.2151 0.5411
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6029 0.1589 10.09 3.7e-11 ***
## FirstAuthorFemale1 0.1123 0.1020 1.10 0.27981
## Year1998 -0.7714 0.1946 -3.96 0.00042 ***
## Year1999 -0.2129 0.1993 -1.07 0.29372
## Year2000 -0.2295 0.3015 -0.76 0.45241
## Year2001 -0.1845 0.2293 -0.80 0.42727
## Year2002 -0.1549 0.1589 -0.97 0.33750
## Year2004 0.0735 0.3322 0.22 0.82640
## Year2005 -0.2915 0.1899 -1.54 0.13519
## Year2006 -0.8078 0.1650 -4.90 3.1e-05 ***
## Year2007 -0.2374 0.2201 -1.08 0.28939
## Year2009 -0.3853 0.1893 -2.04 0.05076 .
```

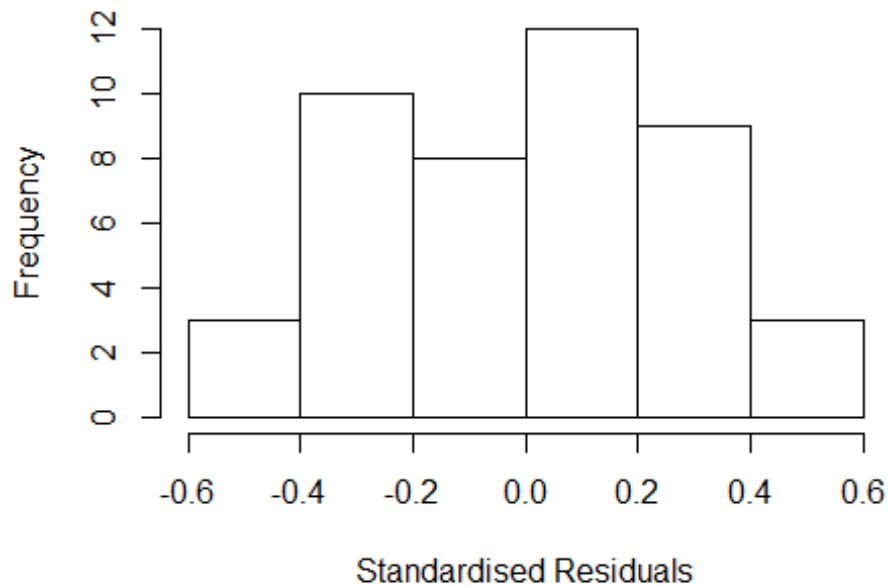


```

## Year2010          -0.0844      0.3198   -0.26   0.79371
## Year2011          -0.7155      0.1650   -4.34   0.00015 ***
## Year2012          -0.6351      0.2522   -2.52   0.01735 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.4
## Multiple R-squared:  0.421, Adjusted R-squared:  0.15
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.820  0.931  0.968  0.951  0.984  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           2.22e-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 1.86 1           1.364
## Year             1.86 13           1.024

```

Residuals from last author



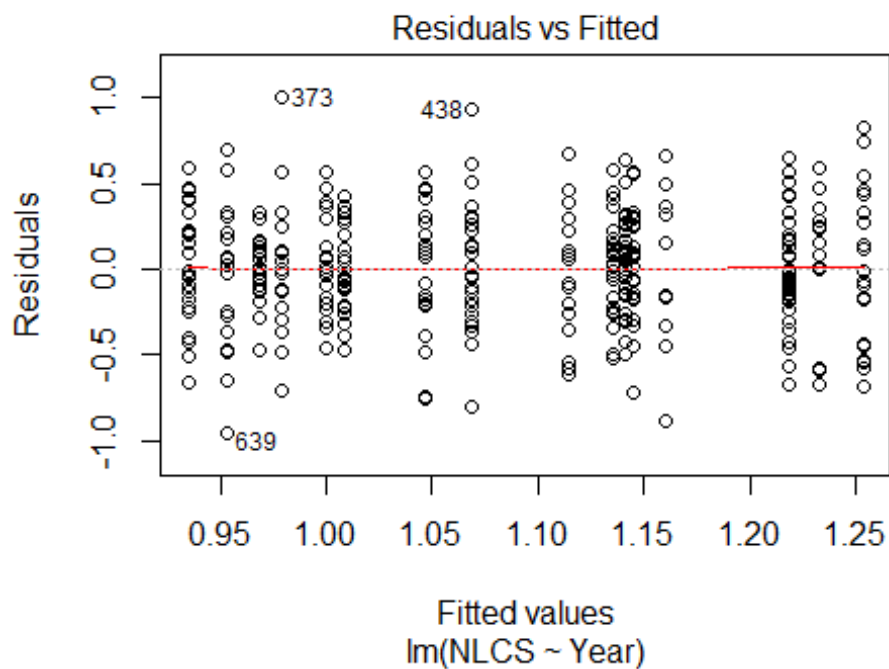
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.606 -0.224 0.009 0.183 0.497
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6469 0.1454 11.33 2.3e-12 ***
## LastAuthorFemale1 -0.0416 0.1289 -0.32 0.74910
## Year1998 -0.8154 0.1839 -4.43 0.00011 ***
## Year1999 -0.2251 0.1926 -1.17 0.25184
## Year2000 -0.2024 0.3068 -0.66 0.51453
## Year2001 -0.1516 0.2856 -0.53 0.59935
## Year2002 -0.1989 0.1454 -1.37 0.18137
## Year2004 0.1064 0.4300 0.25 0.80631
## Year2005 -0.2095 0.1754 -1.19 0.24175
## Year2006 -0.7769 0.1460 -5.32 9.4e-06 ***
## Year2007 -0.2230 0.2159 -1.03 0.30990
## Year2009 -0.3767 0.1839 -2.05 0.04938 *
```

```

## Year2010          -0.1076      0.3016   -0.36   0.72369
## Year2011          -0.7034      0.1459   -4.82   3.9e-05 ***
## Year2012          -0.5461      0.2294   -2.38   0.02386 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.41,   Adjusted R-squared:  0.135
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.787  0.903  0.948  0.938  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      2.22e-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: 45"
## [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
##   30   30   41   35   36   25   21   27   26   22   27   33   15   16   28
## 2011 2012
##   35   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   26   23   34   25   27   21   17   27   20   18   25   29   10   14   21
## 2011 2012
##   23   19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

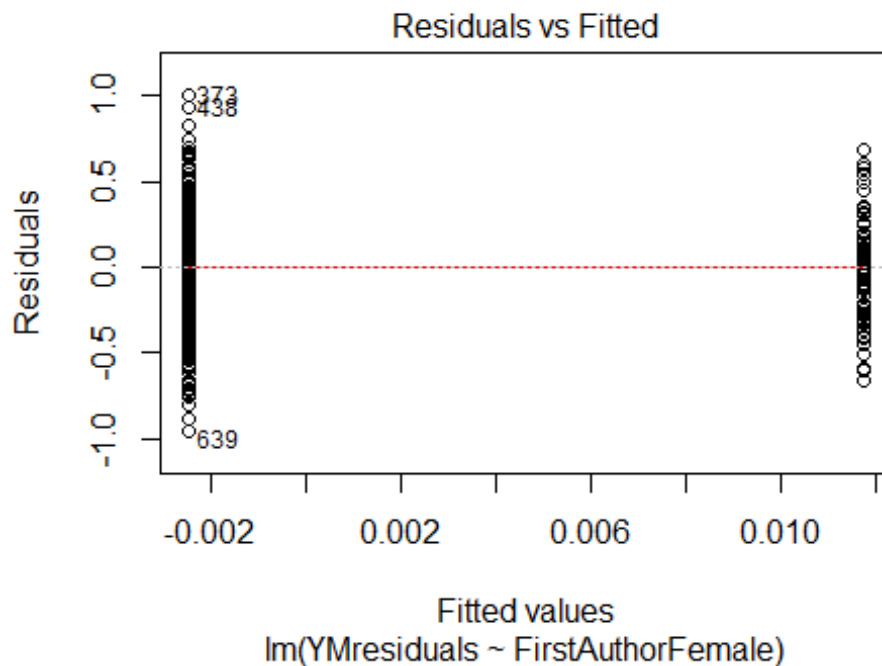
```
##      22      17      30      25      23      18      12      25      17      15      19      25      8      13      18
## 2011 2012
##      19      16
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## 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 = 0.92, df = 1, p-value = 0.3
##
## [1] "Female first author team size 2018 geometric mean: 7.86222418262669"
## [1] "Male first author team size 2018 geometric mean: 2.99379516552391"
##
## 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 = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.66284150148471"
## [1] "Male last author team size 2018 geometric mean: 4.54714969953119"

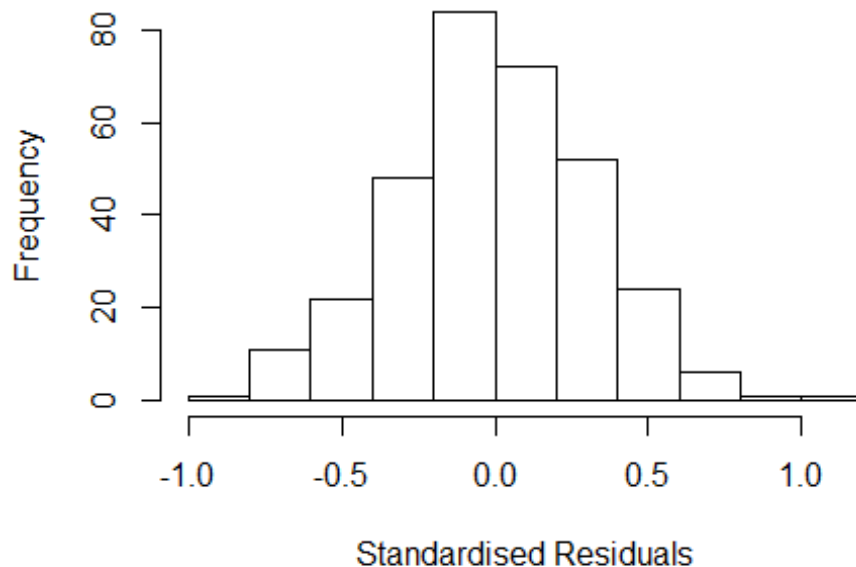
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5.5, 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	1.374	1	1.172
LastAuthorFemale	1.553	1	1.246
UniqueAuthors	2.903	4	1.142
Year	5.087	16	1.052

Residuals from first and last author and team size



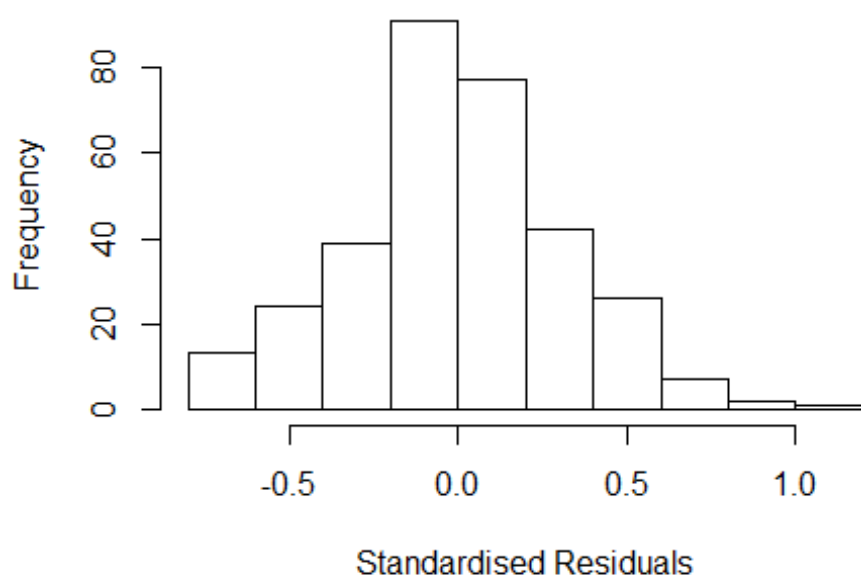
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.83530 -0.20794 -0.00507 0.20856 1.00248
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2423 0.1443 8.61 4.3e-16 ***
## FirstAuthorFemale1 0.0265 0.0505 0.52 0.60057
## LastAuthorFemale1 0.0379 0.0529 0.72 0.47409
## UniqueAuthors2 0.0455 0.1390 0.33 0.74344
## UniqueAuthors3 -0.0989 0.1381 -0.72 0.47412
## UniqueAuthors4 -0.0787 0.1421 -0.55 0.58007
## UniqueAuthors5 -0.0255 0.1453 -0.18 0.86083
## Year1997 -0.0426 0.0970 -0.44 0.66078
## Year1998 -0.0795 0.0628 -1.27 0.20656
## Year1999 -0.0625 0.0824 -0.76 0.44852
```

```

## Year2000          -0.2295      0.0677   -3.39  0.00079 ***
## Year2001          -0.2260      0.0885   -2.55  0.01114 *
## Year2002          -0.1368      0.1182   -1.16  0.24809
## Year2003          -0.1159      0.0735   -1.58  0.11613
## Year2004          -0.2487      0.0669   -3.72  0.00024 ***
## Year2005          -0.2658      0.1066   -2.49  0.01319 *
## Year2006          -0.2555      0.0906   -2.82  0.00512 **
## Year2007          -0.1825      0.0849   -2.15  0.03241 *
## Year2008          -0.1425      0.1819   -0.78  0.43405
## Year2009           0.0686      0.1533    0.45  0.65474
## Year2010          -0.0114      0.1298   -0.09  0.93019
## Year2011          -0.1646      0.0969   -1.70  0.09036 .
## Year2012          -0.1839      0.1395   -1.32  0.18828
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.316
## Multiple R-squared:  0.115, Adjusted R-squared:  0.05
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 289 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.294  0.873  0.954  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          3.11e-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 1.372 1          1.171
## LastAuthorFemale  1.413 1          1.189
## Year              1.929 16          1.021

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -0.770 -0.187 -0.014 0.195 1.035
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21722 0.05228 23.28 < 2e-16 ***
## FirstAuthorFemale1 -0.00393 0.05013 -0.08 0.93764
## LastAuthorFemale1 0.02140 0.05180 0.41 0.67982
## Year1997 -0.03790 0.09720 -0.39 0.69688
## Year1998 -0.07848 0.06849 -1.15 0.25281
## Year1999 -0.06976 0.07842 -0.89 0.37444
## Year2000 -0.24105 0.07003 -3.44 0.00066 ***
## Year2001 -0.25234 0.08577 -2.94 0.00351 **
## Year2002 -0.13609 0.11677 -1.17 0.24477
## Year2003 -0.10347 0.07546 -1.37 0.17135
## Year2004 -0.25042 0.06836 -3.66 0.00029 ***
## Year2005 -0.27300 0.11412 -2.39 0.01736 *
```

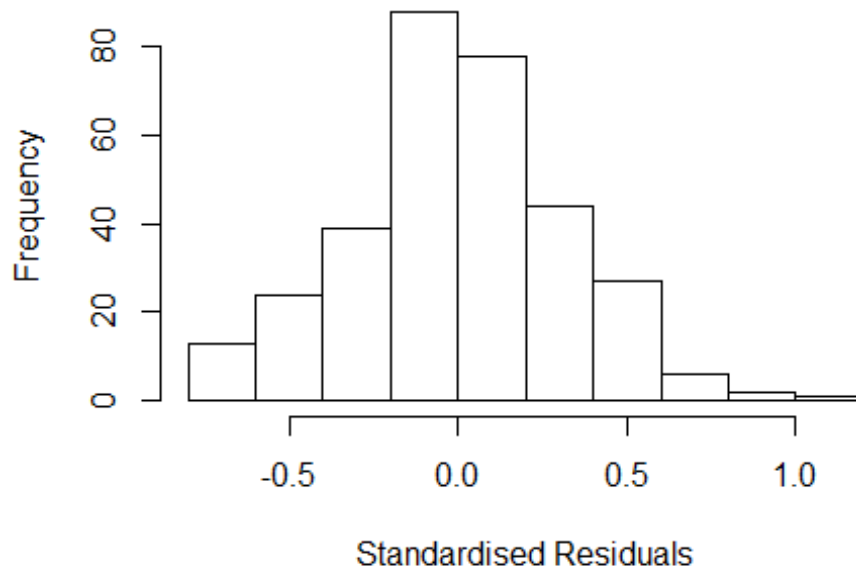


```

## Year2006          -0.27406      0.09028      -3.04      0.00261 **
## Year2007          -0.19669      0.08583      -2.29      0.02261 *
## Year2008          -0.17182      0.17925      -0.96      0.33854
## Year2009           0.04248      0.16336       0.26      0.79499
## Year2010          -0.00789      0.13275      -0.06      0.95266
## Year2011          -0.15769      0.09840      -1.60      0.11007
## Year2012          -0.17521      0.13349      -1.31      0.19033
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.314
## Multiple R-squared:  0.0853, Adjusted R-squared:  0.031
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 296 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##   0.254  0.860  0.957   0.904   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-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 1.376 1      1.173
## Year              1.376 16      1.010

```

Residuals from first author



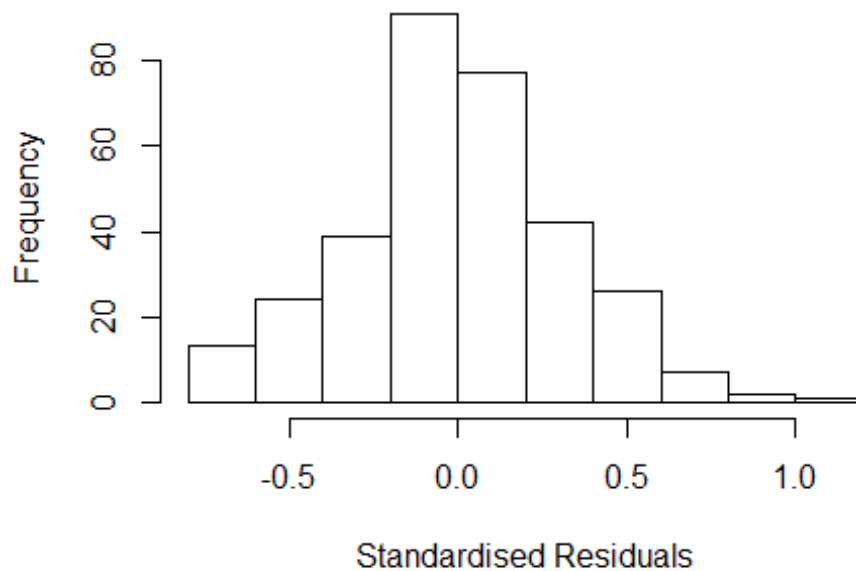
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.77498 -0.18848 -0.00848 0.19946 1.02955
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22208 0.05001 24.44 < 2e-16 ***
## FirstAuthorFemale1 -0.00566 0.05019 -0.11 0.91031
## Year1997 -0.04230 0.09612 -0.44 0.66024
## Year1998 -0.08189 0.06708 -1.22 0.22316
## Year1999 -0.07256 0.07814 -0.93 0.35379
## Year2000 -0.24253 0.06927 -3.50 0.00053 ***
## Year2001 -0.25694 0.08446 -3.04 0.00255 **
## Year2002 -0.13985 0.11631 -1.20 0.23013
## Year2003 -0.10717 0.07428 -1.44 0.15009
## Year2004 -0.24820 0.06801 -3.65 0.00031 ***
## Year2005 -0.27263 0.11373 -2.40 0.01712 *
## Year2006 -0.27333 0.09055 -3.02 0.00275 **
```

```

## Year2007          -0.19700    0.08632   -2.28  0.02316 *
## Year2008          -0.17210    0.17888   -0.96  0.33675
## Year2009           0.04633    0.16194    0.29  0.77501
## Year2010          -0.00311    0.13043   -0.02  0.98098
## Year2011          -0.15141    0.09970   -1.52  0.12987
## Year2012          -0.17012    0.13043   -1.30  0.19311
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.313
## Multiple R-squared:  0.0848, Adjusted R-squared:  0.0336
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 292 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.257  0.857  0.954  0.903  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-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 1.418 1          1.191
## Year              1.418 16          1.011

```

Residuals from last author



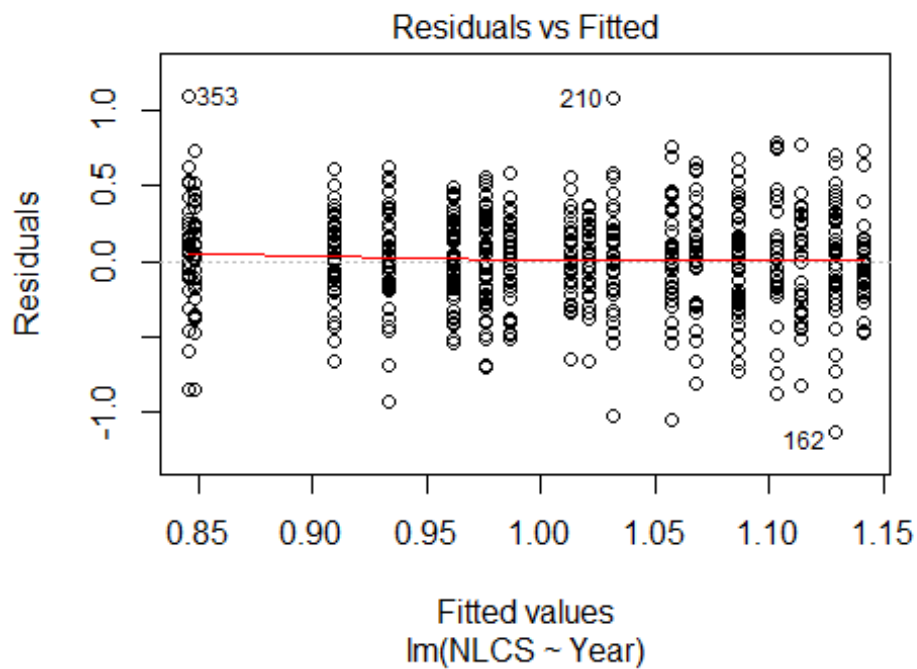
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7703 -0.1871 -0.0146 0.1945 1.0349
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21620 0.05025 24.20 < 2e-16 ***
## LastAuthorFemale1 0.02174 0.05195 0.42 0.67594
## Year1997 -0.03785 0.09722 -0.39 0.69728
## Year1998 -0.07791 0.06810 -1.14 0.25350
## Year1999 -0.06952 0.07814 -0.89 0.37431
## Year2000 -0.24136 0.06973 -3.46 0.00061 ***
## Year2001 -0.25207 0.08584 -2.94 0.00357 **
## Year2002 -0.13519 0.11518 -1.17 0.24142
## Year2003 -0.10383 0.07506 -1.38 0.16760
## Year2004 -0.24975 0.06758 -3.70 0.00026 ***
## Year2005 -0.27211 0.11328 -2.40 0.01690 *
## Year2006 -0.27405 0.09028 -3.04 0.00261 **
```

```

## Year2007          -0.19577      0.08390      -2.33  0.02029 *
## Year2008          -0.17085      0.17880      -0.96  0.34008
## Year2009           0.04287      0.16250       0.26  0.79210
## Year2010          -0.00725      0.13242      -0.05  0.95635
## Year2011          -0.15872      0.09663      -1.64  0.10151
## Year2012          -0.17529      0.13373      -1.31  0.19091
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.314
## Multiple R-squared:  0.0853, Adjusted R-squared:  0.0342
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 297 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.254  0.860  0.957  0.905  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      3.11e-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: 322"
## [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
##   43   59   65   59   47   52   49   43   45   47   56   47   51   62   48
## 2011 2012
##   63   77
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   34   41   46   43   31   35   35   33   37   43   46   33   37   49   34
## 2011 2012

```

```
## 51 53
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 30 32 42 41 26 32 29 30 32 39 39 30 32 45 25
## 2011 2012
## 46 48
## [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
```



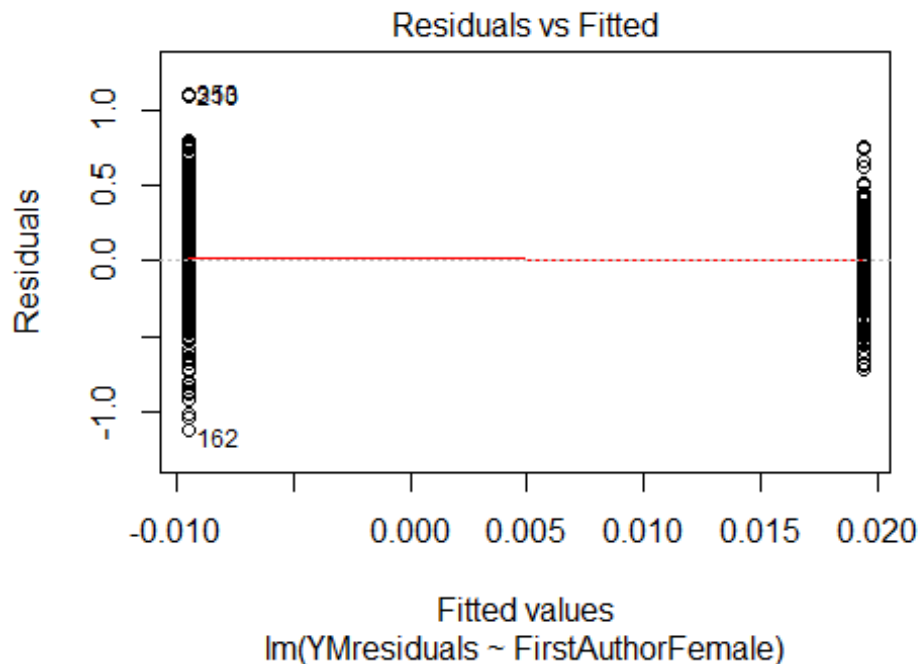
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 24, df = 1, p-value = 1e-06

## [1] "Female first author team size 2018 geometric mean: 4.24264068711929"
## [1] "Male first author team size 2018 geometric mean: 3.63807464040931"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

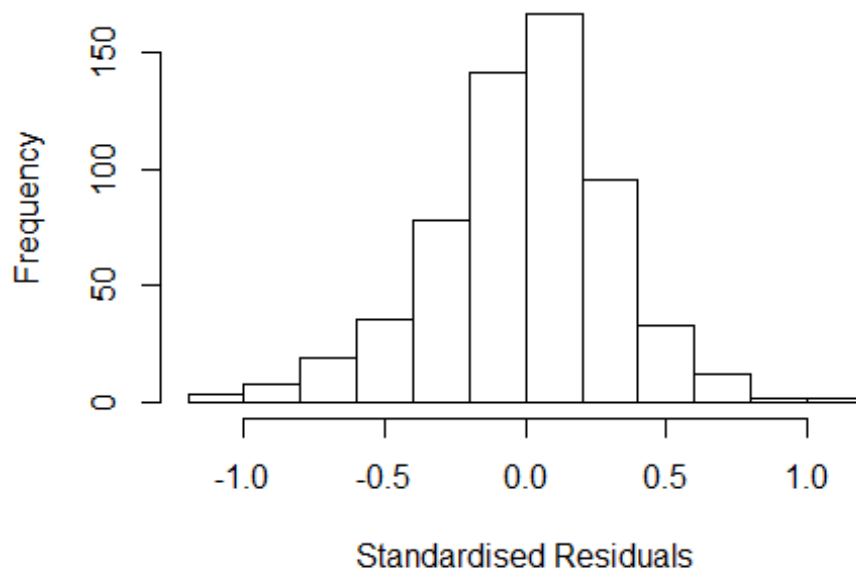
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 57, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.19019970034041"
## [1] "Male last author team size 2018 geometric mean: 3.70608916535354"

## 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 = 0.5
## 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 1.243  1      1.115
## LastAuthorFemale  1.439  1      1.199
## UniqueAuthors    1.773  4      1.074
## Year              2.878 16      1.034
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.1625 -0.1931  0.0135  0.1934  1.1495
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0768    0.1057   10.19 < 2e-16 ***
## FirstAuthorFemale1  0.0340    0.0271    1.25  0.21049
## LastAuthorFemale1 -0.0152    0.0344   -0.44  0.65843
## UniqueAuthors2     0.0550    0.1121    0.49  0.62380
## UniqueAuthors3     0.0178    0.1102    0.16  0.87194
## UniqueAuthors4     0.0533    0.1115    0.48  0.63301
## UniqueAuthors5     0.1229    0.1112    1.11  0.26938
## Year1997          -0.0784    0.0916   -0.86  0.39210
## Year1998           0.0459    0.0719    0.64  0.52346
## Year1999          -0.0926    0.0647   -1.43  0.15265
```

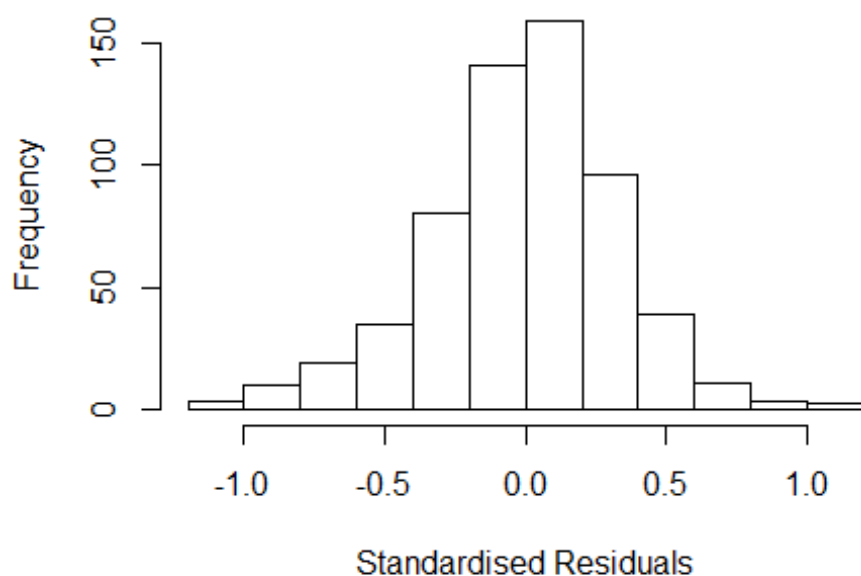


```

## Year2000          -0.0757      0.0914    -0.83   0.40807
## Year2001          -0.2863      0.0831    -3.45   0.00061 ***
## Year2002          -0.1433      0.0605    -2.37   0.01814 *
## Year2003          -0.1116      0.0615    -1.81   0.07039 .
## Year2004          -0.2567      0.0701    -3.66   0.00027 ***
## Year2005          -0.1964      0.0685    -2.87   0.00432 **
## Year2006          -0.2440      0.0633    -3.86   0.00013 ***
## Year2007          -0.1474      0.0583    -2.53   0.01176 *
## Year2008          -0.1165      0.0669    -1.74   0.08209 .
## Year2009          -0.0691      0.0655    -1.06   0.29164
## Year2010          -0.1151      0.0778    -1.48   0.13926
## Year2011          -0.2027      0.0612    -3.31   0.00099 ***
## Year2012          -0.1909      0.0626    -3.05   0.00238 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.304
## Multiple R-squared:  0.0831, Adjusted R-squared:  0.048
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 52 weights are ~= 1. The remaining 546 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.110  0.865   0.955   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          1.67e-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 1.216 1          1.103
## LastAuthorFemale  1.406 1          1.186
## Year              1.663 16          1.016

```

Residuals from first and last author



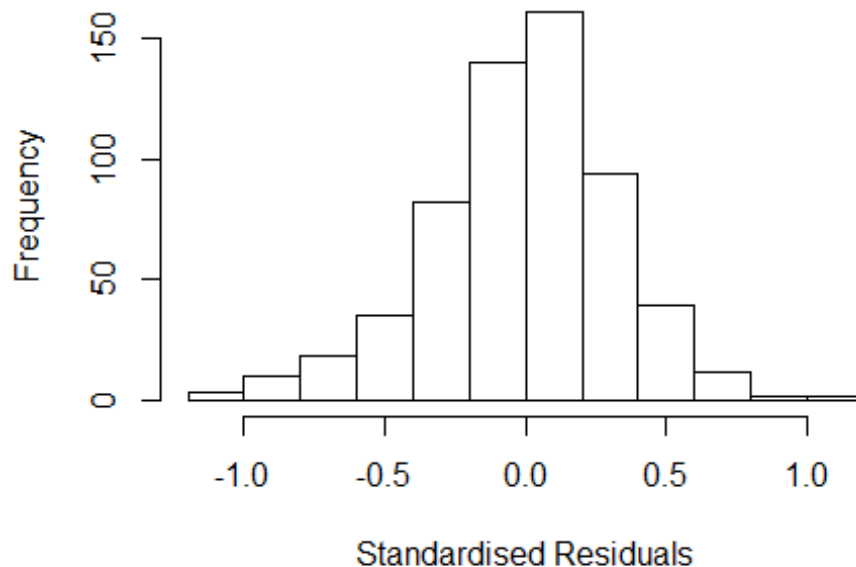
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1556 -0.1965  0.0129  0.2008  1.1002
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1210     0.0426   26.32 < 2e-16 ***
## FirstAuthorFemale1  0.0347     0.0273    1.27  0.20483
## LastAuthorFemale1 -0.0205     0.0346   -0.59  0.55331
## Year1997          -0.0787     0.0912   -0.86  0.38837
## Year1998           0.0552     0.0749    0.74  0.46109
## Year1999          -0.0823     0.0635   -1.30  0.19579
## Year2000          -0.0723     0.0928   -0.78  0.43642
## Year2001          -0.2811     0.0817   -3.44  0.00062 ***
## Year2002          -0.1384     0.0603   -2.29  0.02218 *
## Year2003          -0.1085     0.0602   -1.80  0.07229 .
## Year2004          -0.2527     0.0712   -3.55  0.00042 ***
## Year2005          -0.1688     0.0705   -2.40  0.01692 *
```

```

## Year2006          -0.2265      0.0654   -3.47  0.00057 ***
## Year2007          -0.1283      0.0612   -2.10  0.03651 *
## Year2008          -0.0941      0.0678   -1.39  0.16599
## Year2009          -0.0554      0.0657   -0.84  0.39951
## Year2010          -0.0959      0.0773   -1.24  0.21503
## Year2011          -0.1806      0.0602   -3.00  0.00280 **
## Year2012          -0.1663      0.0624   -2.67  0.00790 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.305
## Multiple R-squared:  0.0685, Adjusted R-squared:  0.0395
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 53 weights are ~= 1. The remaining 545 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.120  0.865   0.953   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      1.67e-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 1.204 1      1.097
## Year              1.204 16      1.006

```

Residuals from first author



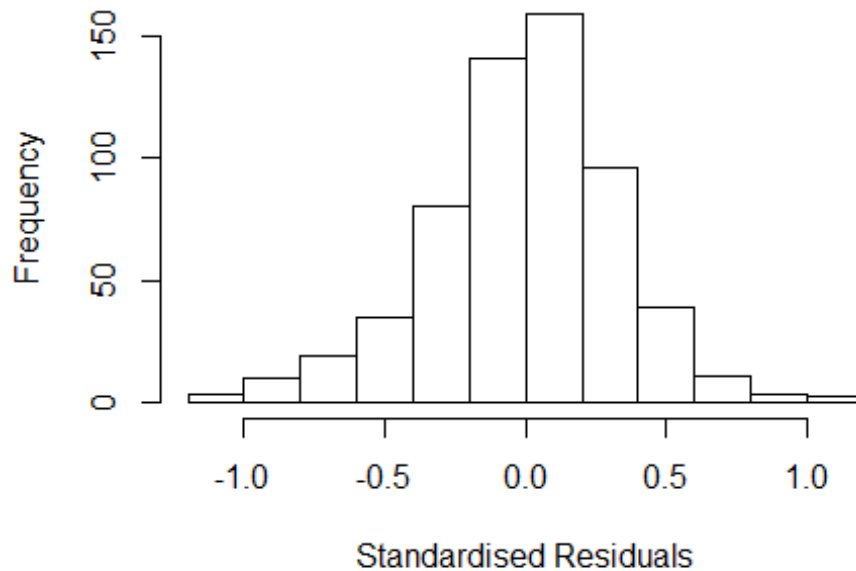
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1722 -0.1976  0.0103  0.1996  1.1013
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1179     0.0419   26.71 < 2e-16 ***
## FirstAuthorFemale1  0.0345     0.0273    1.26  0.20805
## Year1997         -0.0765     0.0904   -0.85  0.39808
## Year1998          0.0543     0.0749    0.73  0.46848
## Year1999         -0.0808     0.0632   -1.28  0.20126
## Year2000         -0.0716     0.0922   -0.78  0.43787
## Year2001         -0.2792     0.0816   -3.42  0.00067 ***
## Year2002         -0.1378     0.0602   -2.29  0.02241 *
## Year2003         -0.1101     0.0603   -1.83  0.06851 .
## Year2004         -0.2528     0.0710   -3.56  0.00041 ***
## Year2005         -0.1688     0.0703   -2.40  0.01673 *
## Year2006         -0.2275     0.0652   -3.49  0.00052 ***
```

```

## Year2007          -0.1356      0.0602   -2.25  0.02473 *
## Year2008          -0.1006      0.0666   -1.51  0.13161
## Year2009          -0.0620      0.0649   -0.96  0.33925
## Year2010          -0.1047      0.0753   -1.39  0.16454
## Year2011          -0.1843      0.0598   -3.08  0.00214 **
## Year2012          -0.1704      0.0619   -2.75  0.00608 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.304
## Multiple R-squared:  0.0675, Adjusted R-squared:  0.0402
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 58 weights are ~= 1. The remaining 540 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.105  0.864  0.953  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      1.67e-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  1.4  1      1.183
## Year              1.4 16      1.011

```

Residuals from last author



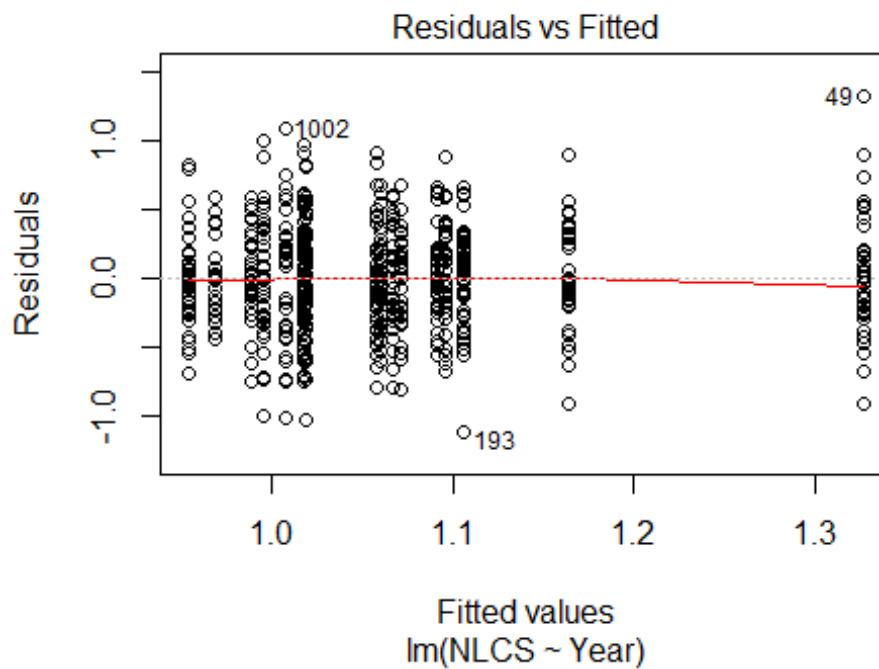
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.15802 -0.19591  0.00891  0.20244  1.09180
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1325     0.0407   27.83 < 2e-16 ***
## LastAuthorFemale1 -0.0200     0.0346   -0.58  0.56248
## Year1997         -0.0803     0.0917   -0.88  0.38122
## Year1998          0.0455     0.0740    0.62  0.53856
## Year1999         -0.0854     0.0636   -1.34  0.17947
## Year2000         -0.0770     0.0915   -0.84  0.40072
## Year2001         -0.2843     0.0808   -3.52  0.00046 ***
## Year2002         -0.1420     0.0603   -2.35  0.01895 *
## Year2003         -0.1065     0.0599   -1.78  0.07566 .
## Year2004         -0.2533     0.0714   -3.55  0.00042 ***
## Year2005         -0.1639     0.0701   -2.34  0.01978 *
## Year2006         -0.2247     0.0655   -3.43  0.00065 ***
```

```

## Year2007          -0.1236      0.0609   -2.03  0.04308 *
## Year2008          -0.0861      0.0681   -1.26  0.20641
## Year2009          -0.0498      0.0654   -0.76  0.44611
## Year2010          -0.0928      0.0767   -1.21  0.22707
## Year2011          -0.1800      0.0601   -3.00  0.00286 **
## Year2012          -0.1588      0.0620   -2.56  0.01070 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.304
## Multiple R-squared:  0.0659, Adjusted R-squared:  0.0385
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 57 weights are ~= 1. The remaining 541 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.116  0.862  0.951   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      1.67e-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: 598"
## [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
##   59   57   66   64   57   42   52   52   59   52   63   61   42   45   54
## 2011 2012
##   64   46
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   44   37   50   37   35   23   33   43   41   39   44   43   32   35   38
## 2011 2012

```

```
## 45 33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 38 29 44 34 31 20 27 37 36 33 36 39 27 32 30
## 2011 2012
## 35 28
## [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
```

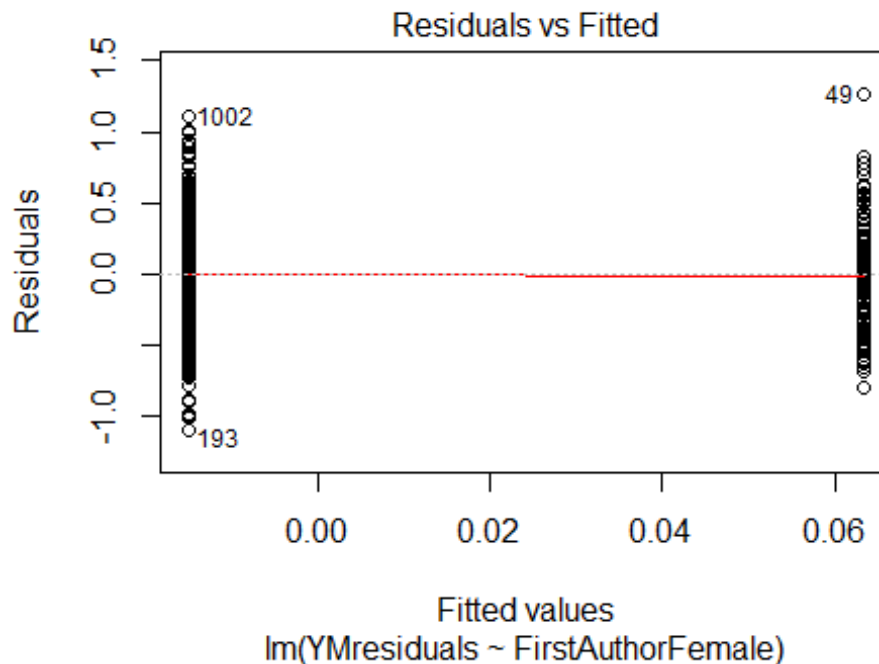


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.41, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 3.96140774326837"
## [1] "Male first author team size 2018 geometric mean: 3.2694873523412"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



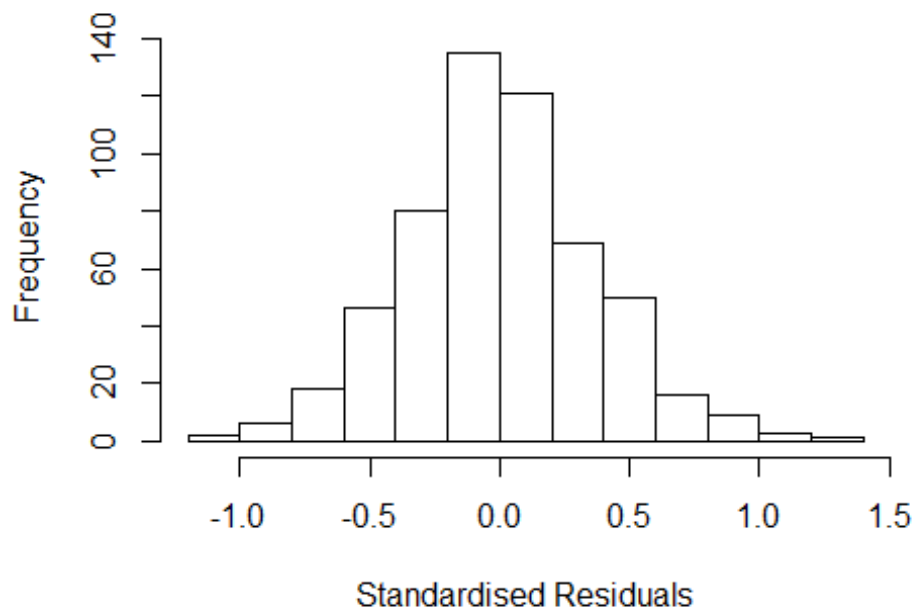
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 72, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.79516302658984"
## [1] "Male last author team size 2018 geometric mean: 3.32771675930544"

## 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 = 70, 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 1.162 1      1.078
## LastAuthorFemale  1.445 1      1.202
## UniqueAuthors    1.571 4      1.058
## Year              2.149 16     1.024
```

Residuals from first and last author and team size



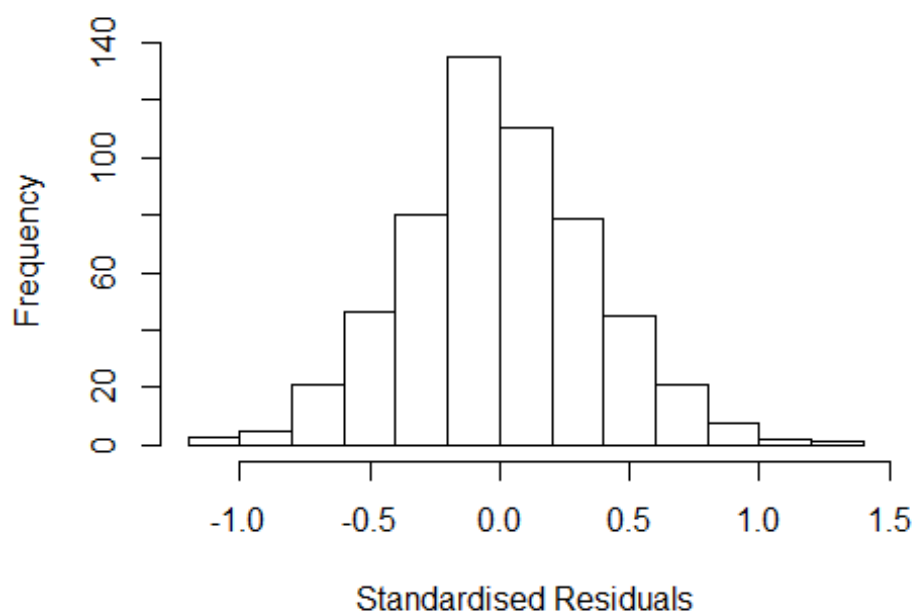
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0378 -0.2259 -0.0093 0.2256 1.2969
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1539 0.0911 12.67 < 2e-16 ***
## FirstAuthorFemale1 0.0712 0.0411 1.73 0.08370 .
## LastAuthorFemale1 0.0322 0.0450 0.71 0.47508
## UniqueAuthors2 0.1311 0.0721 1.82 0.06961 .
## UniqueAuthors3 0.1607 0.0717 2.24 0.02553 *
## UniqueAuthors4 0.1041 0.0751 1.39 0.16602
## UniqueAuthors5 0.1584 0.0793 2.00 0.04617 *
## Year1997 -0.1905 0.0976 -1.95 0.05137 .
## Year1998 -0.1764 0.0810 -2.18 0.02990 *
## Year1999 -0.2148 0.0850 -2.53 0.01183 *
```

```

## Year2000          -0.2579      0.0854   -3.02  0.00265 **
## Year2001          -0.3428      0.0909   -3.77  0.00018 ***
## Year2002          -0.1954      0.0995   -1.96  0.05017 .
## Year2003          -0.2430      0.0790   -3.08  0.00219 **
## Year2004          -0.3404      0.0773   -4.40  1.3e-05 ***
## Year2005          -0.2830      0.0886   -3.19  0.00149 **
## Year2006          -0.3835      0.0800   -4.79  2.1e-06 ***
## Year2007          -0.3113      0.0827   -3.76  0.00019 ***
## Year2008          -0.2347      0.0969   -2.42  0.01574 *
## Year2009          -0.3394      0.1279   -2.65  0.00822 **
## Year2010          -0.2799      0.0980   -2.86  0.00446 **
## Year2011          -0.2767      0.1107   -2.50  0.01276 *
## Year2012          -0.2744      0.1015   -2.70  0.00712 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.0769, Adjusted R-squared:  0.0388
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 44 weights are ~= 1. The remaining 512 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.144  0.860  0.954  0.901  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          1.80e-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 1.139 1          1.067
## LastAuthorFemale  1.384 1          1.176
## Year              1.573 16          1.014

```

Residuals from first and last author



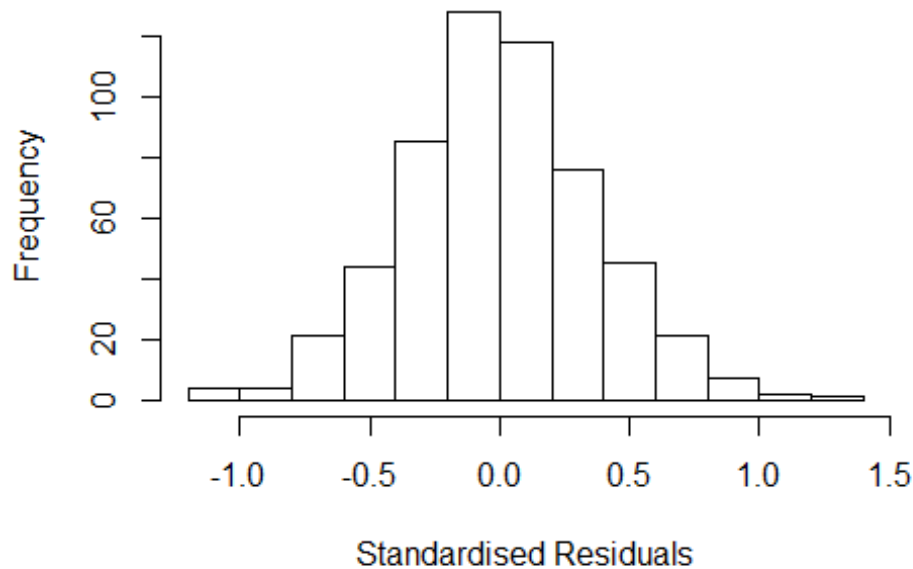
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1046 -0.2299 -0.0149  0.2286  1.2928
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2755     0.0636   20.06 < 2e-16 ***
## FirstAuthorFemale1  0.0847     0.0407    2.08  0.03780 *
## LastAuthorFemale1  0.0376     0.0444    0.85  0.39831
## Year1997          -0.1830     0.0984   -1.86  0.06353 .
## Year1998          -0.1709     0.0819   -2.09  0.03738 *
## Year1999          -0.2106     0.0869   -2.42  0.01572 *
## Year2000          -0.2610     0.0855   -3.05  0.00237 **
## Year2001          -0.3495     0.0900   -3.88  0.00012 ***
## Year2002          -0.1919     0.0970   -1.98  0.04851 *
## Year2003          -0.2393     0.0788   -3.04  0.00252 **
## Year2004          -0.3288     0.0773   -4.25  2.5e-05 ***
## Year2005          -0.2787     0.0894   -3.12  0.00192 **
```

```

## Year2006          -0.3823      0.0794   -4.81  1.9e-06 ***
## Year2007          -0.3057      0.0828   -3.69  0.00025 ***
## Year2008          -0.2517      0.0963   -2.61  0.00921 **
## Year2009          -0.3327      0.1311   -2.54  0.01143 *
## Year2010          -0.2848      0.0997   -2.86  0.00447 **
## Year2011          -0.2662      0.1113   -2.39  0.01717 *
## Year2012          -0.2704      0.0977   -2.77  0.00586 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.0633, Adjusted R-squared:  0.0319
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 46 weights are ~= 1. The remaining 510 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.148  0.859  0.953  0.899  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      1.80e-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 1.141 1      1.068
## Year              1.141 16      1.004

```

Residuals from first author



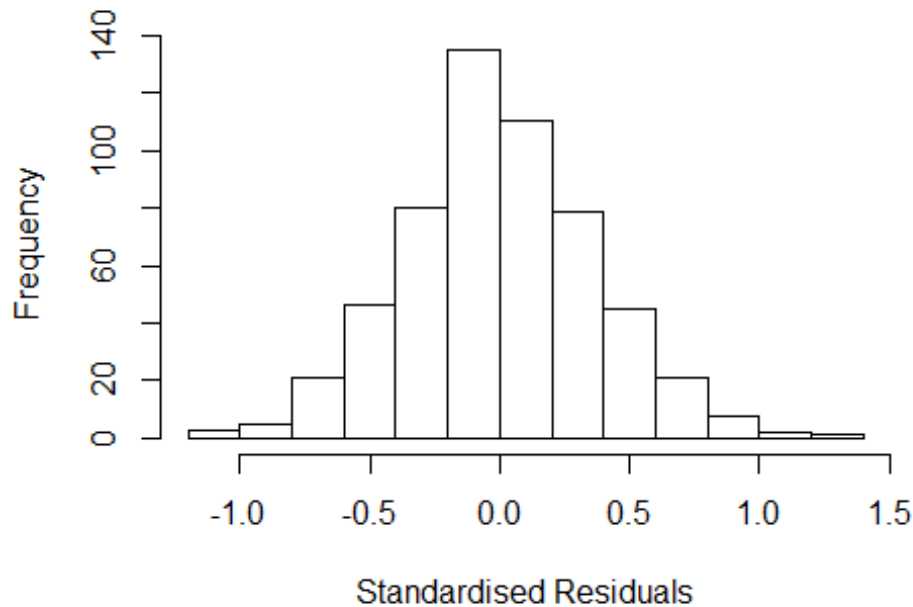
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.10642 -0.22332 -0.00971 0.23506 1.28580
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2822 0.0623 20.59 < 2e-16 ***
## FirstAuthorFemale1 0.0850 0.0407 2.09 0.03703 *
## Year1997 -0.1887 0.0981 -1.92 0.05495 .
## Year1998 -0.1758 0.0811 -2.17 0.03054 *
## Year1999 -0.2151 0.0864 -2.49 0.01309 *
## Year2000 -0.2640 0.0850 -3.11 0.00200 **
## Year2001 -0.3562 0.0891 -4.00 7.3e-05 ***
## Year2002 -0.1958 0.0969 -2.02 0.04383 *
## Year2003 -0.2448 0.0781 -3.13 0.00182 **
## Year2004 -0.3279 0.0774 -4.24 2.7e-05 ***
## Year2005 -0.2797 0.0888 -3.15 0.00173 **
## Year2006 -0.3805 0.0794 -4.79 2.2e-06 ***
```

```

## Year2007          -0.3060      0.0832   -3.68  0.00026 ***
## Year2008          -0.2547      0.0964   -2.64  0.00844 **
## Year2009          -0.3214      0.1287   -2.50  0.01280 *
## Year2010          -0.2792      0.0993   -2.81  0.00512 **
## Year2011          -0.2558      0.1110   -2.30  0.02159 *
## Year2012          -0.2704      0.0974   -2.77  0.00571 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.062, Adjusted R-squared:  0.0323
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 48 weights are ~= 1. The remaining 508 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.154  0.854  0.952  0.899  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      1.80e-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 1.376 1          1.173
## Year              1.376 16          1.010

```

Residuals from last author



```
## [1] "List of 0 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
## -1.11217 -0.23174 -0.00795 0.23139 1.36002
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2930 0.0611 21.17 < 2e-16 ***
## LastAuthorFemale1 0.0384 0.0445 0.86 0.3877
## Year1997 -0.1746 0.0963 -1.81 0.0704 .
## Year1998 -0.1808 0.0797 -2.27 0.0237 *
## Year1999 -0.2112 0.0858 -2.46 0.0141 *
## Year2000 -0.2606 0.0833 -3.13 0.0019 **
## Year2001 -0.3622 0.0879 -4.12 4.4e-05 ***
## Year2002 -0.2035 0.0955 -2.13 0.0336 *
## Year2003 -0.2365 0.0777 -3.04 0.0024 **
## Year2004 -0.3290 0.0767 -4.29 2.2e-05 ***
## Year2005 -0.2700 0.0920 -2.93 0.0035 **
## Year2006 -0.3749 0.0794 -4.72 3.0e-06 ***
```

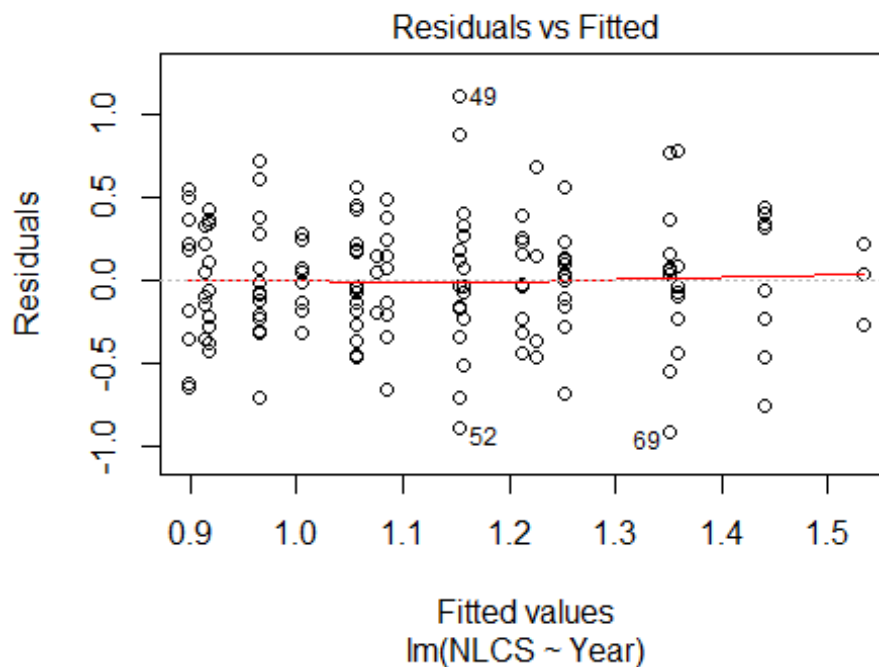


```

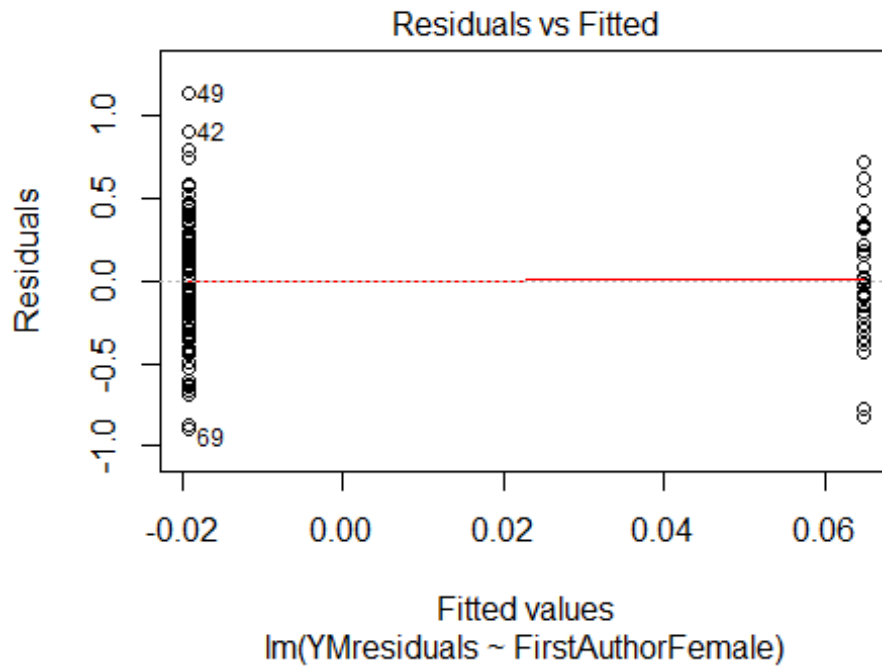
## Year2007          -0.3145      0.0804    -3.91    0.0001 ***
## Year2008          -0.2589      0.0935    -2.77    0.0058 **
## Year2009          -0.3340      0.1328    -2.51    0.0122 *
## Year2010          -0.2899      0.0991    -2.93    0.0036 **
## Year2011          -0.2555      0.1103    -2.32    0.0209 *
## Year2012          -0.2622      0.0986    -2.66    0.0080 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.354
## Multiple R-squared:  0.0557, Adjusted R-squared:  0.0258
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 43 weights are ~= 1. The remaining 513 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.106  0.864  0.952  0.900  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      1.80e-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: 556"
## [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
##    8   12   6   16   11    5   10   13   13   11   18   15   12   19   13
## 2011 2012
##    9   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    8    4   10    8    3    9   10    7    9   14   12    9   15    9
## 2011 2012

```

```
##      6      8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      3      6      3      8      6      3      8      9      6      8      12      9      8      12      5
## 2011 2012
##      4      7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



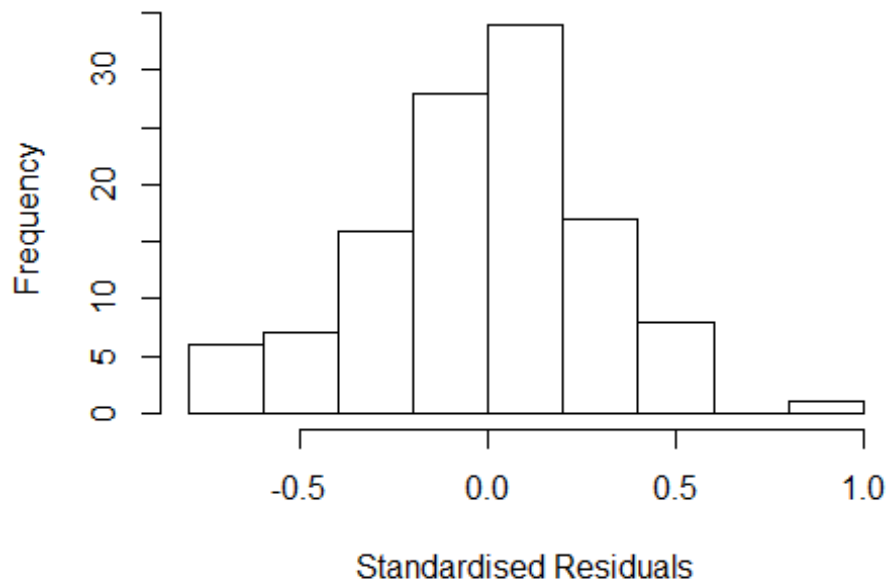
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.068, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 4"
## [1] "Male first author team size 2018 geometric mean: 2"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 2.37841423000544"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF^(1/(2*Df))
## FirstAuthorFemale	2.187	1	1.479
## LastAuthorFemale	2.442	1	1.563
## UniqueAuthors	10.394	4	1.340
## Year	32.497	16	1.115

Residuals from first and last author and team size



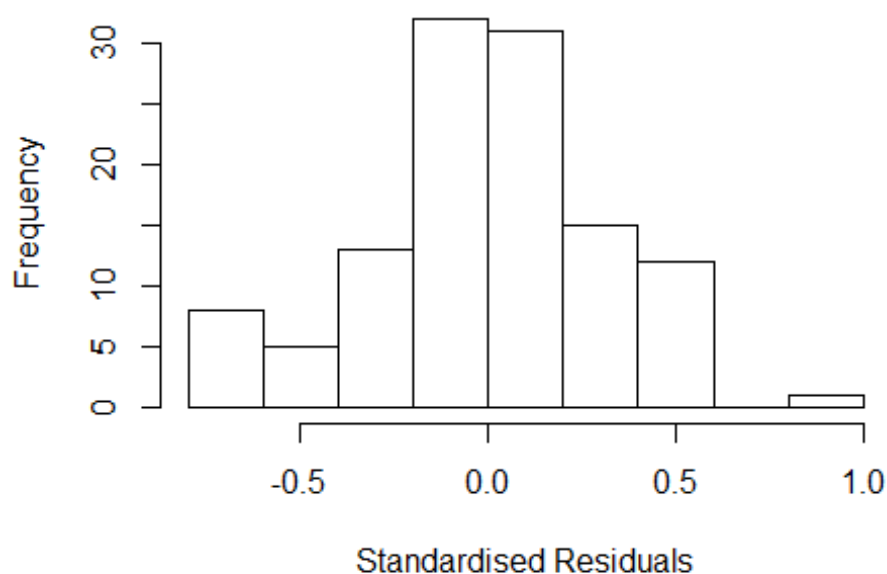
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.74076 -0.19790 0.00306 0.17910 0.82420
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7852 0.1639 4.79 6.2e-06 ***
## FirstAuthorFemale1 0.1071 0.0899 1.19 0.2368
## LastAuthorFemale1 0.2969 0.1046 2.84 0.0055 **
## UniqueAuthors2 0.2922 0.1406 2.08 0.0404 *
## UniqueAuthors3 0.2874 0.1568 1.83 0.0699 .
## UniqueAuthors4 0.3395 0.1464 2.32 0.0226 *
## UniqueAuthors5 0.2781 0.1507 1.85 0.0682 .
## Year1997 0.2408 0.2621 0.92 0.3605
## Year1998 -0.1920 0.1387 -1.38 0.1697
## Year1999 -0.1414 0.1885 -0.75 0.4552
```

```

## Year2000          0.1635      0.2266      0.72      0.4723
## Year2001          0.3248      0.1301      2.50      0.0143 *
## Year2002         -0.1776      0.3013     -0.59      0.5571
## Year2003         -0.1688      0.1351     -1.25      0.2145
## Year2004          0.2008      0.1498      1.34      0.1835
## Year2005         -0.1084      0.1392     -0.78      0.4380
## Year2006         -0.2599      0.1430     -1.82      0.0723 .
## Year2007          0.0232      0.1380      0.17      0.8667
## Year2008          0.0139      0.1347      0.10      0.9180
## Year2009         -0.1577      0.1216     -1.30      0.1981
## Year2010          0.0627      0.1788      0.35      0.7268
## Year2011         -0.3195      0.1306     -2.45      0.0163 *
## Year2012         -0.2185      0.1316     -1.66      0.1002
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.296
## Multiple R-squared:  0.367, Adjusted R-squared:  0.219
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 107 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.418  0.866  0.958   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      8.55e-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 1.978 1      1.406
## LastAuthorFemale  1.767 1      1.329
## Year              2.970 16      1.035

```

Residuals from first and last author



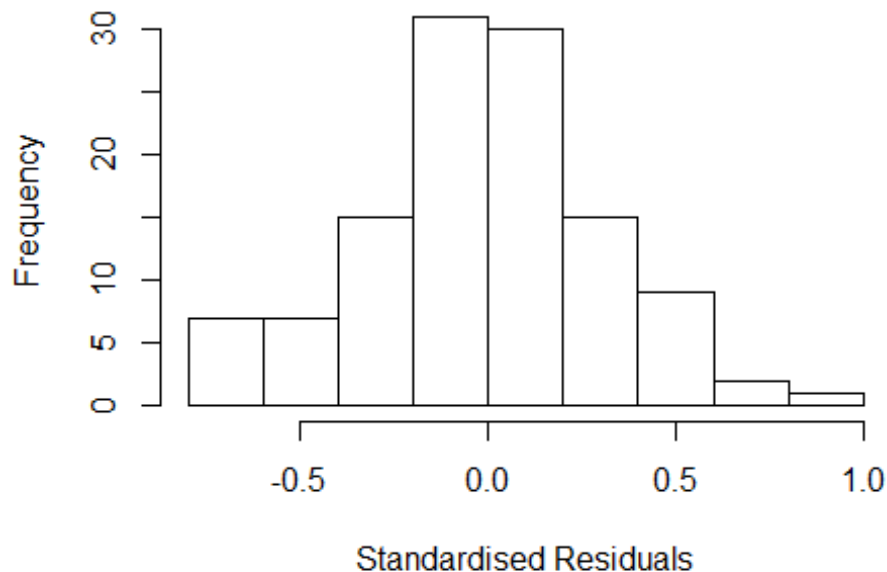
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7434 -0.1854 0.0129 0.1892 0.8935
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0757 0.0873 12.32 <2e-16 ***
## FirstAuthorFemale1 0.1169 0.0850 1.37 0.1725
## LastAuthorFemale1 0.2851 0.0951 3.00 0.0035 **
## Year1997 0.2358 0.2358 1.00 0.3197
## Year1998 -0.1760 0.1223 -1.44 0.1532
## Year1999 -0.1414 0.1784 -0.79 0.4302
## Year2000 0.1005 0.2519 0.40 0.6907
## Year2001 0.3225 0.1241 2.60 0.0108 *
## Year2002 -0.1864 0.2813 -0.66 0.5090
## Year2003 -0.2044 0.1424 -1.44 0.1542
## Year2004 0.1766 0.1414 1.25 0.2148
## Year2005 -0.1327 0.1583 -0.84 0.4039
```

```

## Year2006          -0.2802      0.1293    -2.17    0.0327 *
## Year2007           0.0339      0.1374     0.25    0.8057
## Year2008          -0.0222      0.1229    -0.18    0.8572
## Year2009          -0.1576      0.1125    -1.40    0.1644
## Year2010           0.0595      0.1684     0.35    0.7246
## Year2011          -0.3278      0.1215    -2.70    0.0082 **
## Year2012          -0.2180      0.1241    -1.76    0.0821 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.303
## Multiple R-squared:  0.335, Adjusted R-squared:  0.213
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.365  0.866  0.959  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      8.55e-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 1.699 1      1.303
## Year              1.699 16      1.017

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.79097 -0.19847 -0.00548 0.19550 0.91262
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.07547 0.08677 12.39 <2e-16 ***
## FirstAuthorFemale1 0.16946 0.08486 2.00 0.0486 *
## Year1997 0.22000 0.22301 0.99 0.3263
## Year1998 -0.09132 0.19237 -0.47 0.6361
## Year1999 -0.15343 0.17852 -0.86 0.3922
## Year2000 0.15550 0.20706 0.75 0.4544
## Year2001 0.40324 0.12676 3.18 0.0020 **
## Year2002 -0.19197 0.25137 -0.76 0.4469
## Year2003 -0.20242 0.13946 -1.45 0.1498
## Year2004 0.16822 0.15541 1.08 0.2817
## Year2005 -0.09710 0.15176 -0.64 0.5238
## Year2006 -0.29909 0.13000 -2.30 0.0235 *
```

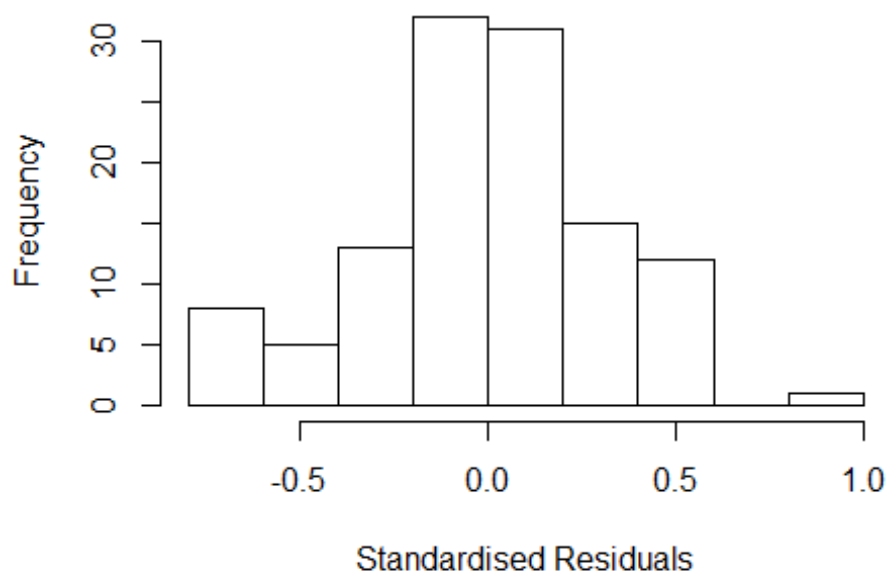


```

## Year2007          0.15042    0.15026    1.00    0.3192
## Year2008          0.00725    0.13386    0.05    0.9569
## Year2009         -0.08699    0.12307   -0.71    0.4814
## Year2010          0.11538    0.14235    0.81    0.4196
## Year2011         -0.34097    0.12414   -2.75    0.0072 **
## Year2012         -0.15985    0.11914   -1.34    0.1828
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.264, Adjusted R-squared:  0.138
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 107 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.439  0.882  0.964  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      8.55e-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 1.742 1          1.320
## Year            1.742 16          1.017

```

Residuals from last author



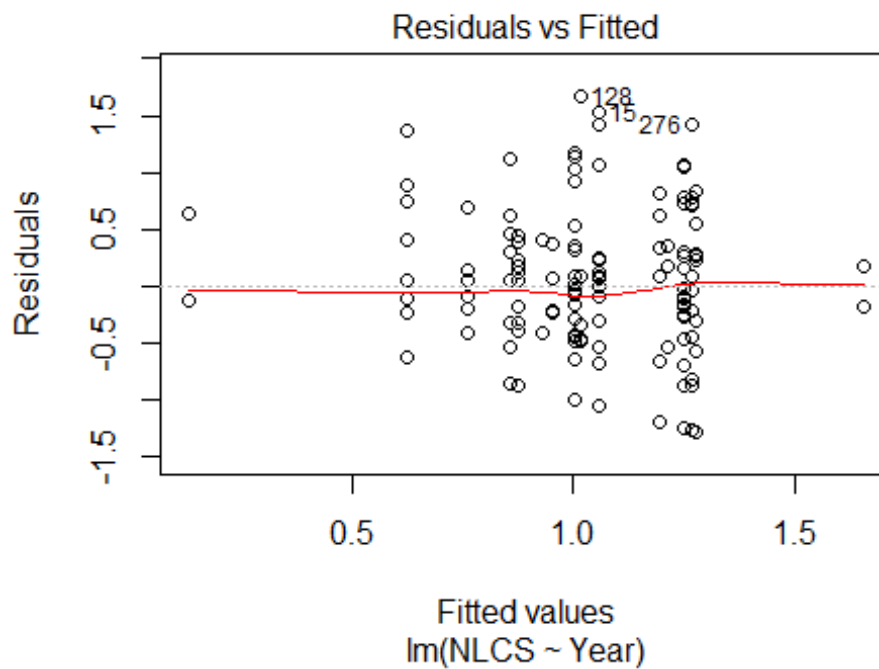
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7687 -0.1774 0.0149 0.1814 0.8350
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.07578 0.08740 12.31 < 2e-16 ***
## LastAuthorFemale1 0.31696 0.09168 3.46 0.00081 ***
## Year1997 0.24511 0.24817 0.99 0.32571
## Year1998 -0.18624 0.11688 -1.59 0.11423
## Year1999 -0.13522 0.16878 -0.80 0.42497
## Year2000 0.13294 0.25631 0.52 0.60514
## Year2001 0.35213 0.12068 2.92 0.00436 **
## Year2002 -0.18537 0.28714 -0.65 0.52005
## Year2003 -0.20478 0.14285 -1.43 0.15485
## Year2004 0.22905 0.14740 1.55 0.12339
## Year2005 -0.08453 0.17060 -0.50 0.62136
## Year2006 -0.22178 0.12237 -1.81 0.07295 .
```

```

## Year2007      0.04771    0.12958    0.37  0.71352
## Year2008     -0.00966    0.12218   -0.08  0.93714
## Year2009     -0.15383    0.11247   -1.37  0.17448
## Year2010      0.06902    0.19217    0.36  0.72024
## Year2011     -0.29694    0.11433   -2.60  0.01083 *
## Year2012     -0.16482    0.12026   -1.37  0.17361
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.299
## Multiple R-squared:  0.322, Adjusted R-squared:  0.206
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 105 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.414  0.852  0.960  0.897  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      8.55e-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: 117"
## [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
##   11  24   6   3  10  14   7   7   5  17  13   8  15  23  26
## 2011 2012
##   20   6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11  16   4   2   3   9   6   7   2   6  10   7  13  21  12
## 2011 2012

```

```
## 13 6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 12 4 2 1 8 6 7 2 6 10 6 13 21 11
## 2011 2012
## 12 4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 19, df = 16, p-value = 0.2
```



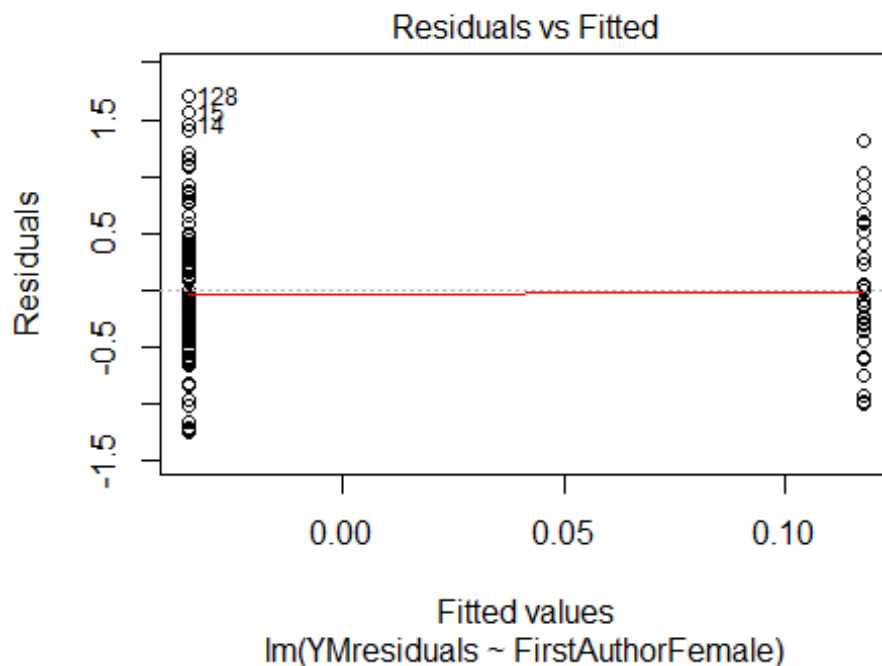
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0011, df = 1, p-value = 1

## [1] "Female first author team size 2018 geometric mean: 5"
## [1] "Male first author team size 2018 geometric mean: 2.29440538087975"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

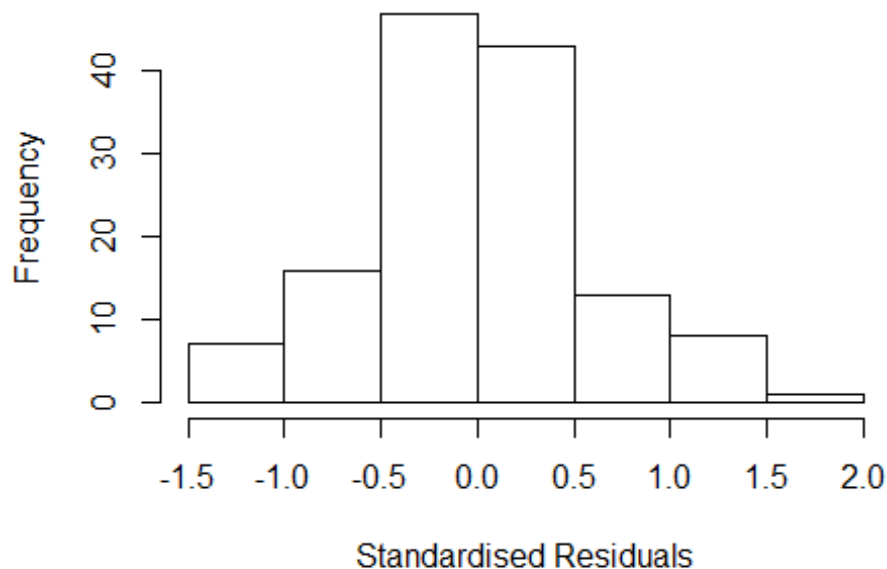
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 2.29440538087975"

## 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 = 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  3.143  1      1.773
## LastAuthorFemale  2.542  1      1.594
## UniqueAuthors    10.723  4      1.345
## Year              27.176 16      1.109
```

Residuals from first and last author and team size



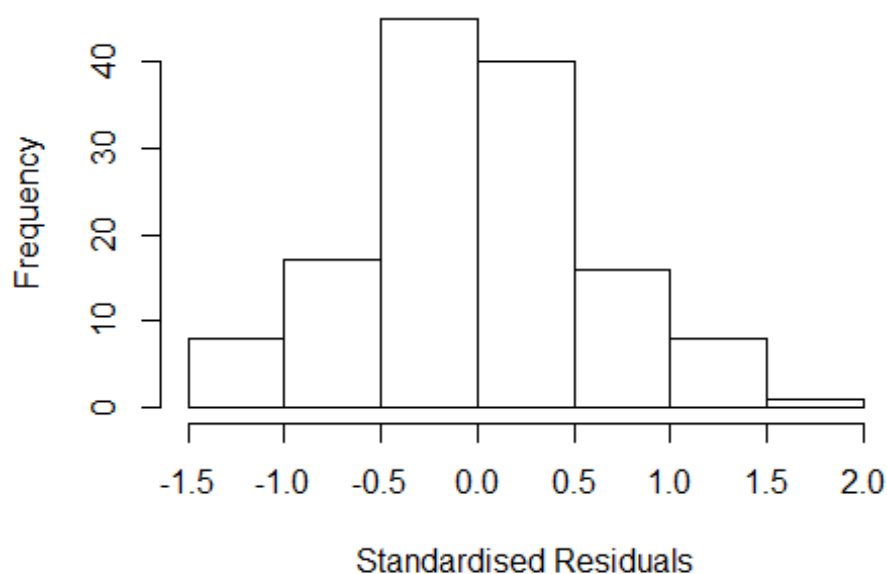
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.29757 -0.35325 -0.00719 0.36722 1.80179
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1853 0.2907 4.08 8.6e-05 ***
## FirstAuthorFemale1 0.3041 0.1489 2.04 0.04344 *
## LastAuthorFemale1 -0.1478 0.1807 -0.82 0.41510
## UniqueAuthors2 0.1412 0.1447 0.98 0.33129
## UniqueAuthors3 0.2725 0.1941 1.40 0.16307
## UniqueAuthors4 0.1888 0.3040 0.62 0.53593
## UniqueAuthors5 0.2472 0.2595 0.95 0.34285
## Year1997 -0.2409 0.3638 -0.66 0.50921
## Year1998 -0.2334 0.3166 -0.74 0.46246
## Year1999 0.3921 0.3621 1.08 0.28122
```

```

## Year2000          -0.5103      0.2907    -1.76   0.08198 .
## Year2001          -0.6671      0.3879    -1.72   0.08828 .
## Year2002          -0.0319      0.4585    -0.07   0.94467
## Year2003          -0.5285      0.2991    -1.77   0.07992 .
## Year2004          -0.4569      0.5216    -0.88   0.38293
## Year2005          -0.4453      0.3875    -1.15   0.25288
## Year2006          -0.0289      0.3406    -0.08   0.93244
## Year2007          -0.1046      0.4188    -0.25   0.80316
## Year2008          -0.5269      0.3033    -1.74   0.08512 .
## Year2009          -0.4161      0.3216    -1.29   0.19831
## Year2010          -0.7757      0.3555    -2.18   0.03119 *
## Year2011          -0.0577      0.3640    -0.16   0.87437
## Year2012          -1.1781      0.3040    -3.87   0.00018 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.582
## Multiple R-squared:  0.234, Adjusted R-squared:  0.0831
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.318  0.865  0.958  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          7.41e-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 2.028 1          1.424
## LastAuthorFemale 2.063 1          1.436
## Year              2.250 16          1.026

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.3257 -0.3653 -0.0408 0.3694 1.9217
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2235 0.2763 4.43 2.2e-05 ***
## FirstAuthorFemale1 0.3374 0.1497 2.25 0.0261 *
## LastAuthorFemale1 -0.1800 0.1802 -1.00 0.3200
## Year1997 -0.2081 0.3547 -0.59 0.5585
## Year1998 -0.2718 0.3034 -0.90 0.3722
## Year1999 0.3533 0.3508 1.01 0.3159
## Year2000 -0.5485 0.2763 -1.98 0.0495 *
## Year2001 -0.5287 0.3810 -1.39 0.1678
## Year2002 0.0535 0.4533 0.12 0.9063
## Year2003 -0.5127 0.2895 -1.77 0.0791 .
## Year2004 -0.2065 0.5099 -0.40 0.6863
## Year2005 -0.4622 0.3645 -1.27 0.2074
```

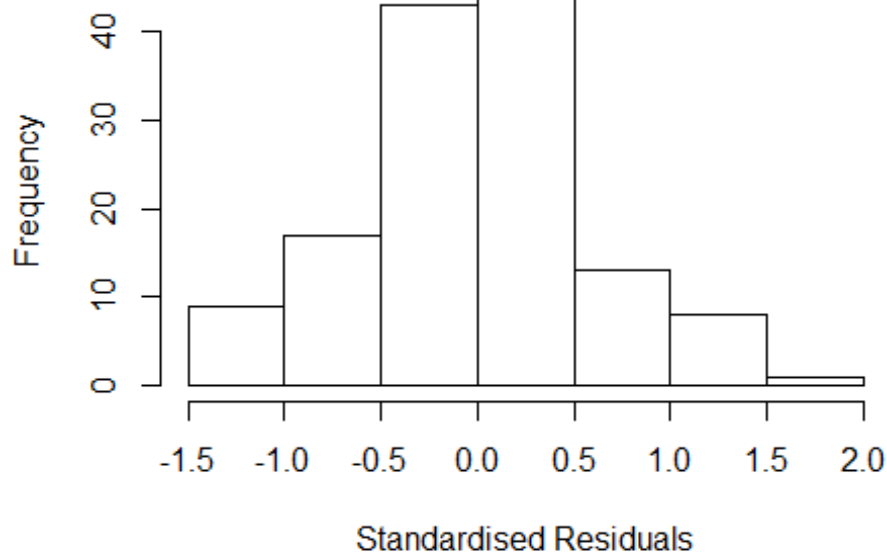


```

## Year2006          0.1022      0.3369      0.30      0.7622
## Year2007          -0.0340      0.3670     -0.09      0.9263
## Year2008          -0.4251      0.2976     -1.43      0.1559
## Year2009          -0.3460      0.3147     -1.10      0.2739
## Year2010          -0.7080      0.3815     -1.86      0.0660 .
## Year2011           0.0138      0.3719      0.04      0.9705
## Year2012          -1.1593      0.3016     -3.84      0.0002 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.555
## Multiple R-squared:  0.228, Adjusted R-squared:  0.108
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 129 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.206  0.862  0.955   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      7.41e-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 1.51 1          1.229
## Year              1.51 16          1.013

```

Residuals from first author



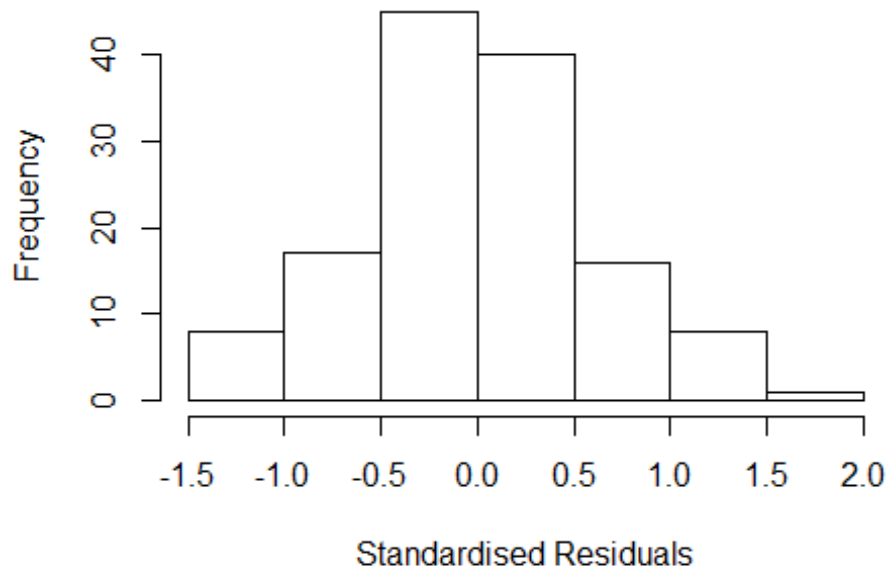
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3184 -0.3609 -0.0122 0.3701 1.9512
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.226961 0.271719 4.52 1.5e-05 ***
## FirstAuthorFemale1 0.268608 0.138566 1.94 0.05497 .
## Year1997 -0.222696 0.352033 -0.63 0.52823
## Year1998 -0.275029 0.299119 -0.92 0.35975
## Year1999 0.294235 0.362541 0.81 0.41867
## Year2000 -0.551961 0.271719 -2.03 0.04448 *
## Year2001 -0.549829 0.356541 -1.54 0.12575
## Year2002 0.035818 0.439669 0.08 0.93521
## Year2003 -0.507243 0.288901 -1.76 0.08174 .
## Year2004 -0.299961 0.412081 -0.73 0.46812
## Year2005 -0.495186 0.347434 -1.43 0.15674
## Year2006 0.091465 0.330854 0.28 0.78269
```

```

## Year2007          -0.015435    0.364484    -0.04    0.96629
## Year2008          -0.436377    0.292053    -1.49    0.13782
## Year2009          -0.369477    0.308965    -1.20    0.23417
## Year2010          -0.686818    0.372714    -1.84    0.06790 .
## Year2011          -0.000889    0.358789     0.00    0.99803
## Year2012          -1.174061    0.307907    -3.81    0.00022 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.594
## Multiple R-squared:  0.211, Adjusted R-squared:  0.0964
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.259  0.872  0.958  0.906  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      7.41e-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 2.096 1      1.448
## Year            2.096 16      1.023

```

Residuals from last author



```
## [1] "List of 0 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
## -1.346126 -0.392758 -0.000247 0.379935 1.877466
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.27924 0.24248 5.28 6.2e-07 ***
## LastAuthorFemale1 -0.01143 0.16428 -0.07 0.94466
## Year1997 -0.27225 0.33024 -0.82 0.41138
## Year1998 -0.32724 0.27276 -1.20 0.23266
## Year1999 0.38198 0.28867 1.32 0.18834
## Year2000 -0.60424 0.24248 -2.49 0.01410 *
## Year2001 -0.55997 0.32944 -1.70 0.09183 .
## Year2002 -0.01972 0.42174 -0.05 0.96278
## Year2003 -0.52884 0.27286 -1.94 0.05502 .
## Year2004 -0.34652 0.40962 -0.85 0.39929
## Year2005 -0.47370 0.33890 -1.40 0.16482
## Year2006 0.03745 0.30792 0.12 0.90340
```

```

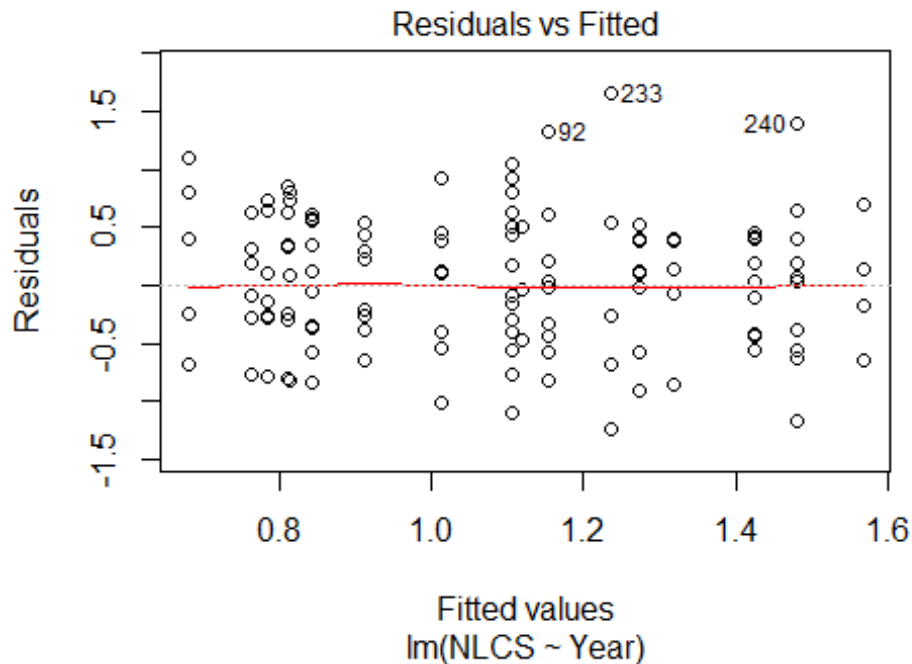
## Year2007          -0.00897      0.34519   -0.03  0.97931
## Year2008          -0.39268      0.26265   -1.50  0.13759
## Year2009          -0.30899      0.29771   -1.04  0.30146
## Year2010          -0.70034      0.34805   -2.01  0.04650 *
## Year2011           0.06689      0.37682     0.18  0.85942
## Year2012          -1.09666      0.30147   -3.64  0.00041 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.185, Adjusted R-squared:  0.0663
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.315  0.875  0.959  0.906  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      7.41e-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: 135"
## [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]"
##
## 2000 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    1    2    2    1    4    3    1    4    2    1
##
## 2000 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    0    2    2    1    3    1    1    3    2    0
##
## 2000 2002 2004 2005 2006 2007 2008 2009 2011 2012
##    0    2    1    1    3    1    1    3    2    0

```

```

## [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: 1"
## [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: 14"
## [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
##    9   13   15   13   14   12   12    9    9   11    9   17   22   15   13
## 2011 2012
##   16   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3    8    7    5    8    9    8    6    7    4   10   10    8    5
## 2011 2012
##   10   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3    6    5    5    8    9    4    6    7    3    8    9    6    4
## 2011 2012
##    7   13
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.7

```



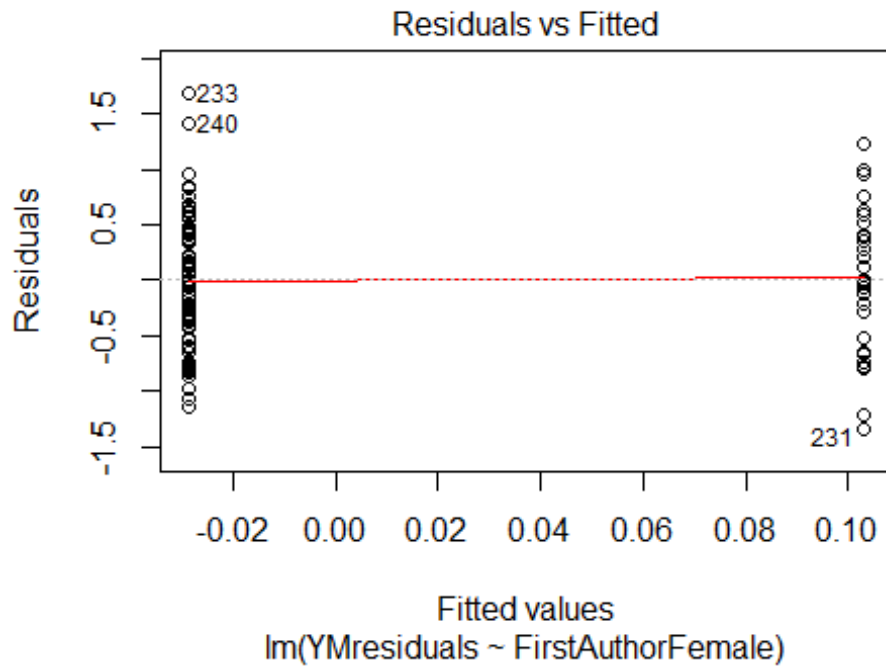
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 1"
## [1] "Male first author team size 2018 geometric mean: 2.28942848510666"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

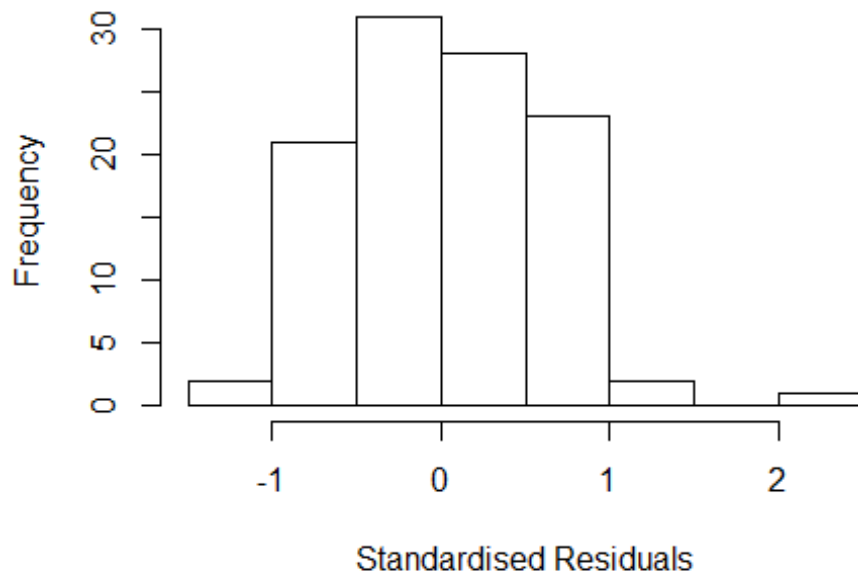
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male last author 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 = 0.5, 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 6.502 1      2.550
## LastAuthorFemale  5.151 1      2.269
## UniqueAuthors    18.920 3      1.632
## Year             54.020 16      1.133
```


Residuals from first and last author and team size



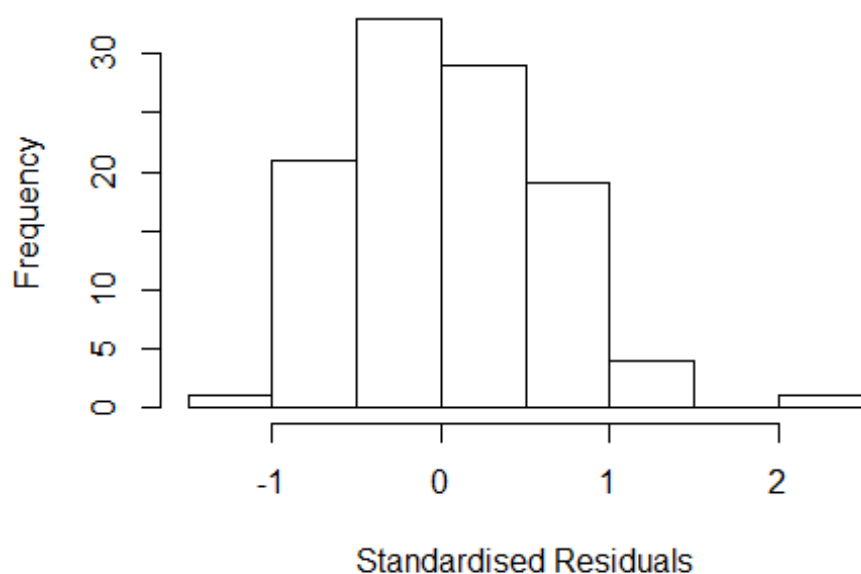
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0673 -0.3791 0.0147 0.4844 2.1026
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.72048 0.42807 1.68 0.096 .
## FirstAuthorFemale1 0.25193 0.29834 0.84 0.401
## LastAuthorFemale1 0.09979 0.23200 0.43 0.668
## UniqueAuthors2 0.25137 0.15050 1.67 0.099 .
## UniqueAuthors3 0.27359 0.16353 1.67 0.098 .
## UniqueAuthors4 0.41177 0.64288 0.64 0.524
## Year1997 0.17799 0.48988 0.36 0.717
## Year1998 0.49078 0.44345 1.11 0.271
## Year1999 0.09358 0.44097 0.21 0.832
## Year2000 0.49132 0.48761 1.01 0.316
```

```

## Year2001      -0.01763    0.46273   -0.04    0.970
## Year2002      0.25621    0.44465    0.58    0.566
## Year2003     -0.07903    0.61294   -0.13    0.898
## Year2004     -0.13932    0.47848   -0.29    0.772
## Year2005     -0.33895    0.57002   -0.59    0.554
## Year2006      0.41397    0.57575    0.72    0.474
## Year2007      0.47747    0.43253    1.10    0.273
## Year2008     -0.06570    0.46905   -0.14    0.889
## Year2009     -0.04874    0.48791   -0.10    0.921
## Year2010     -0.34489    1.11471   -0.31    0.758
## Year2011      0.37049    0.46077    0.80    0.424
## Year2012     -0.00488    0.47898   -0.01    0.992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.629
## Multiple R-squared:  0.226, Adjusted R-squared:  0.0367
## Convergence in 37 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 102 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.240  0.899  0.952  0.928  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      9.26e-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 3.212 1 1.792
## LastAuthorFemale 3.023 1 1.739
## Year 2.588 16 1.030

```

Residuals from first and last author



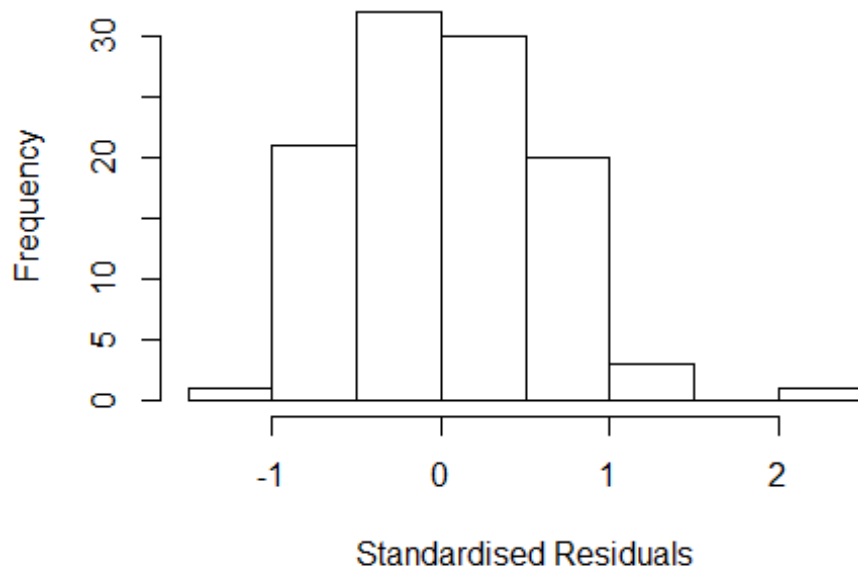
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.236 -0.372 -0.029 0.427 2.464
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7296 0.4089 1.78 0.078 .
## FirstAuthorFemale1 0.2364 0.2148 1.10 0.274
## LastAuthorFemale1 0.0964 0.1869 0.52 0.607
## Year1997 0.3117 0.4413 0.71 0.482
## Year1998 0.5780 0.4216 1.37 0.174
## Year1999 0.1906 0.4428 0.43 0.668
## Year2000 0.5896 0.4649 1.27 0.208
## Year2001 0.1015 0.4324 0.23 0.815
## Year2002 0.3270 0.4440 0.74 0.463
## Year2003 0.1738 0.5776 0.30 0.764
## Year2004 -0.0214 0.4373 -0.05 0.961
## Year2005 -0.1440 0.5000 -0.29 0.774
```

```

## Year2006          0.5574      0.4731      1.18      0.242
## Year2007          0.5822      0.4145      1.40      0.164
## Year2008          0.0738      0.4447      0.17      0.869
## Year2009          0.0327      0.4911      0.07      0.947
## Year2010         -0.3038      0.6791     -0.45      0.656
## Year2011          0.3945      0.4583      0.86      0.392
## Year2012          0.1736      0.4407      0.39      0.695
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.672
## Multiple R-squared:  0.191, Adjusted R-squared:  0.0274
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 98 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.149  0.902  0.954  0.930  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      9.26e-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 1.839 1      1.356
## Year              1.839 16      1.019

```

Residuals from first author



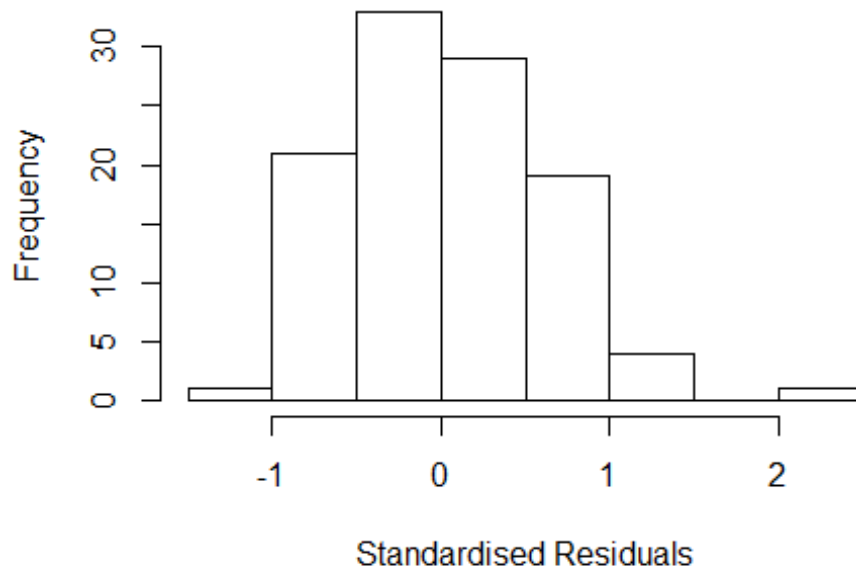
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.234974 -0.365851  0.000377  0.418655  2.496182
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7322     0.4159   1.76   0.082 .
## FirstAuthorFemale1 0.3173     0.1620   1.96   0.053 .
## Year1997          0.2826     0.4335   0.65   0.516
## Year1998          0.5784     0.4285   1.35   0.180
## Year1999          0.1713     0.4415   0.39   0.699
## Year2000          0.6106     0.4690   1.30   0.196
## Year2001          0.1081     0.4409   0.25   0.807
## Year2002          0.3251     0.4509   0.72   0.473
## Year2003          0.1696     0.5861   0.29   0.773
## Year2004         -0.0210     0.4440  -0.05   0.962
## Year2005         -0.1552     0.5151  -0.30   0.764
## Year2006          0.5639     0.4822   1.17   0.245
```

```

## Year2007          0.5981      0.4229      1.41      0.161
## Year2008          0.0850      0.4504      0.19      0.851
## Year2009          0.0337      0.5001      0.07      0.946
## Year2010         -0.3384      0.6939     -0.49      0.627
## Year2011          0.3986      0.4646      0.86      0.393
## Year2012          0.1855      0.4490      0.41      0.680
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.648
## Multiple R-squared:  0.195, Adjusted R-squared:  0.0428
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 98 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.104  0.894  0.951  0.925  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      9.26e-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 1.542 1      1.242
## Year              1.542 16      1.014

```

Residuals from last author



```
## [1] "List of 0 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
## -1.1721 -0.4538 -0.0417 0.4527 2.3633
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7478 0.3940 1.90 0.061 .
## LastAuthorFemale1 0.2713 0.1406 1.93 0.057 .
## Year1997 0.3691 0.4624 0.80 0.427
## Year1998 0.5704 0.4100 1.39 0.168
## Year1999 0.2184 0.4457 0.49 0.625
## Year2000 0.5315 0.4487 1.18 0.239
## Year2001 0.0937 0.4175 0.22 0.823
## Year2002 0.3214 0.4370 0.74 0.464
## Year2003 0.1567 0.5660 0.28 0.782
## Year2004 -0.0284 0.4270 -0.07 0.947
## Year2005 -0.1229 0.4840 -0.25 0.800
## Year2006 0.5804 0.4761 1.22 0.226
```

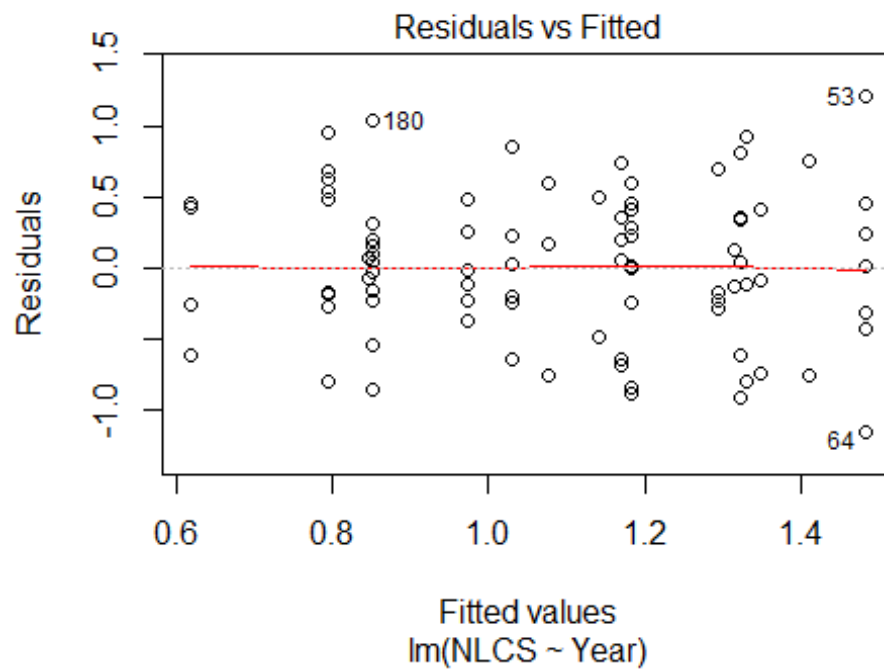
```

## Year2007          0.5629      0.4011      1.40      0.164
## Year2008          0.0438      0.4365      0.10      0.920
## Year2009          0.0240      0.4817      0.05      0.960
## Year2010         -0.2211      0.7350     -0.30      0.764
## Year2011          0.4075      0.4502      0.91      0.368
## Year2012          0.1530      0.4288      0.36      0.722
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.693
## Multiple R-squared:  0.175, Adjusted R-squared:  0.0195
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 100 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.221  0.917  0.958  0.935  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      9.26e-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: 108"
## [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
##    9    4    5    3    3    5    7   12   10   15    6   10    8   17   15
## 2011 2012
##   12   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    2    3    2    2    3    2    7    6   12    4    6    4   11   10
## 2011 2012

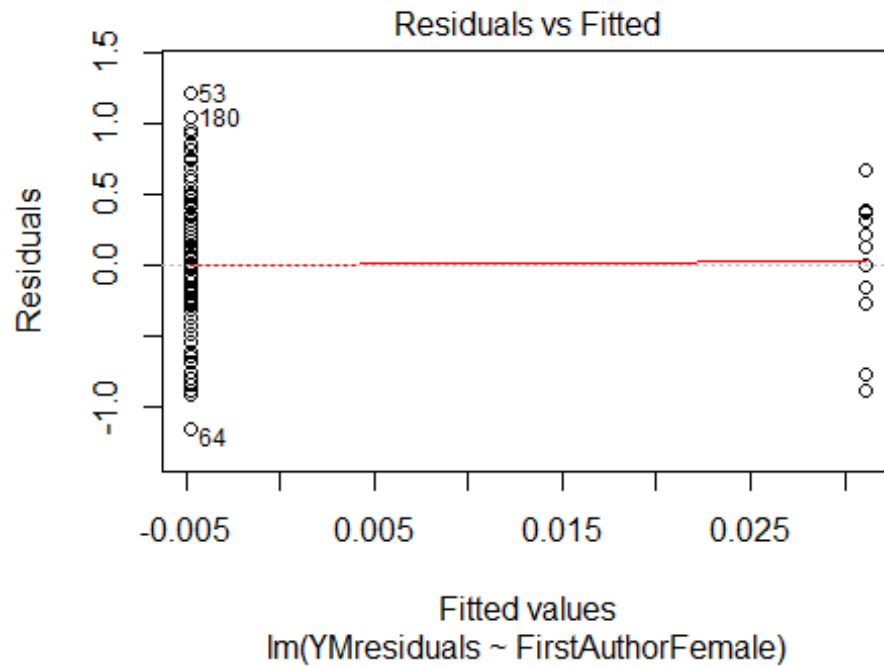
```



```
##      6      6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    2    2    2    2    3    2    6    4   10    4    6    4    8   10
## 2011 2012
##    5    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 9.8, df = 16, p-value = 0.9
```

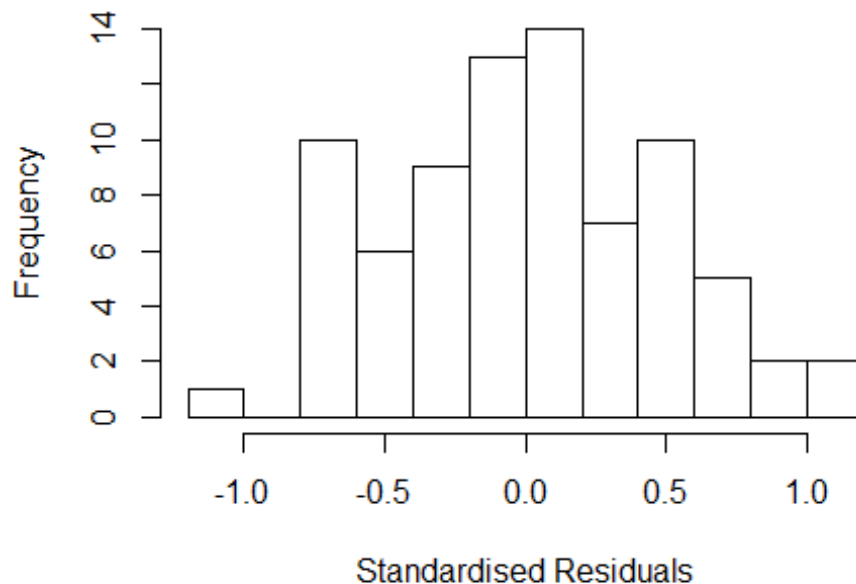


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.37, df = 1, p-value = 0.5
```



```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 1.41421356237309"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  32.28  1          5.682
## LastAuthorFemale   50.40  1          7.099
## UniqueAuthors    1543.96  4          2.504
## Year              9504.69 16          1.331
```

Residuals from first and last author and team size



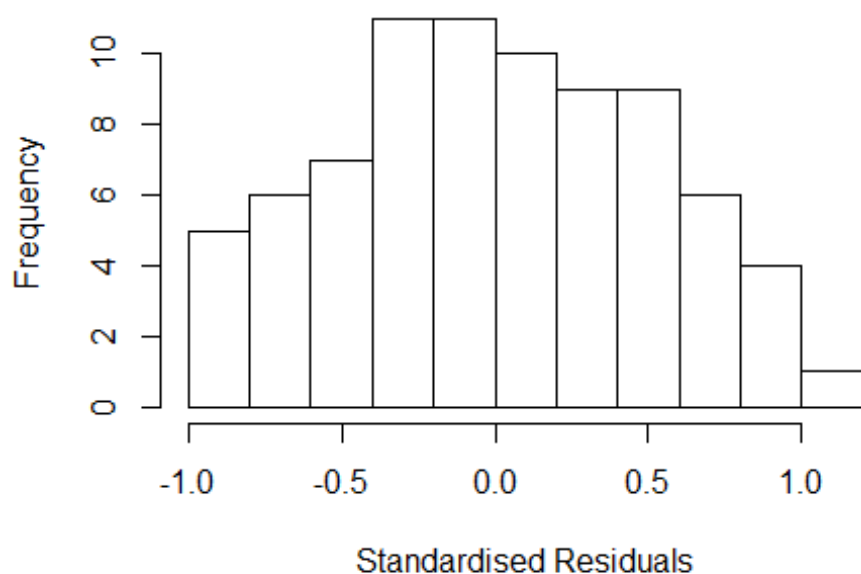
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06175 -0.31722 0.00806 0.32501 1.09628
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6243 0.2513 2.48 0.01600 *
## FirstAuthorFemale1 -0.3278 0.2102 -1.56 0.12449
## LastAuthorFemale1 0.4049 0.2775 1.46 0.15011
## UniqueAuthors2 0.2787 0.1424 1.96 0.05529 .
## UniqueAuthors3 0.2986 0.1955 1.53 0.13244
## UniqueAuthors4 1.1082 0.2831 3.91 0.00025 ***
## UniqueAuthors5 -0.0927 0.2426 -0.38 0.70379
## Year1997 0.6463 0.6091 1.06 0.29317
## Year1998 0.0193 0.3600 0.05 0.95738
## Year1999 0.5162 0.4680 1.10 0.27479
```

```

## Year2000          0.5128      0.2964      1.73  0.08907 .
## Year2001          0.4838      0.4973      0.97  0.33479
## Year2002          0.0838      0.2667      0.31  0.75448
## Year2003          0.9504      0.3280      2.90  0.00536 **
## Year2004          0.2817      0.4070      0.69  0.49172
## Year2005          0.1036      0.3293      0.31  0.75424
## Year2006          0.5298      0.3594      1.47  0.14602
## Year2007          0.0507      0.3154      0.16  0.87282
## Year2008          0.4068      0.2741      1.48  0.14334
## Year2009          0.1614      0.3525      0.46  0.64886
## Year2010          0.4329      0.2995      1.45  0.15390
## Year2011          0.1449      0.2951      0.49  0.62534
## Year2012          0.4539      0.4195      1.08  0.28387
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.601
## Multiple R-squared:  0.302, Adjusted R-squared:  0.0277
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.720  0.908  0.960  0.942  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          1.27e-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 2.653 1          1.629
## LastAuthorFemale 2.645 1          1.626
## Year              3.166 16          1.037

```

Residuals from first and last author



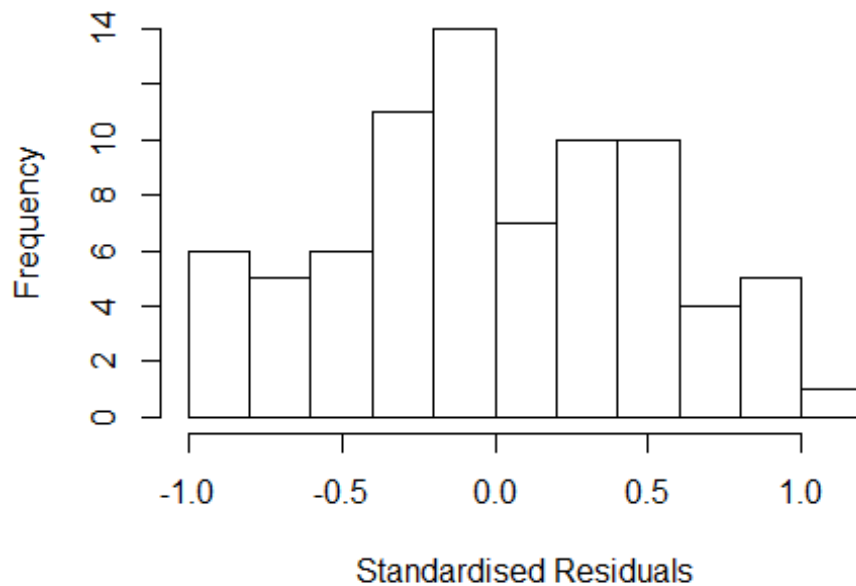
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.87997 -0.29676 -0.00161 0.39906 1.04396
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.62401 0.25009 2.50 0.0154 *
## FirstAuthorFemale1 0.00111 0.22359 0.00 0.9961
## LastAuthorFemale1 0.28346 0.24485 1.16 0.2516
## Year1997 0.78599 0.80297 0.98 0.3316
## Year1998 0.15899 0.44874 0.35 0.7244
## Year1999 0.51649 0.46437 1.11 0.2705
## Year2000 0.54871 0.33803 1.62 0.1098
## Year2001 0.67529 0.56900 1.19 0.2400
## Year2002 0.22349 0.25490 0.88 0.3841
## Year2003 1.01903 0.35602 2.86 0.0058 **
## Year2004 0.45815 0.33490 1.37 0.1764
## Year2005 0.19614 0.33010 0.59 0.5546
```

```

## Year2006          0.74030      0.34381      2.15      0.0353 *
## Year2007          0.30028      0.28768      1.04      0.3008
## Year2008          0.65076      0.31733      2.05      0.0447 *
## Year2009          0.29225      0.29706      0.98      0.3292
## Year2010          0.54996      0.29775      1.85      0.0697 .
## Year2011          0.34167      0.31242      1.09      0.2785
## Year2012          0.62856      0.41205      1.53      0.1324
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.203, Adjusted R-squared:  -0.0357
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.755  0.913   0.961   0.939   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      1.27e-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 1.97 1      1.404
## Year              1.97 16      1.021

```

Residuals from first author



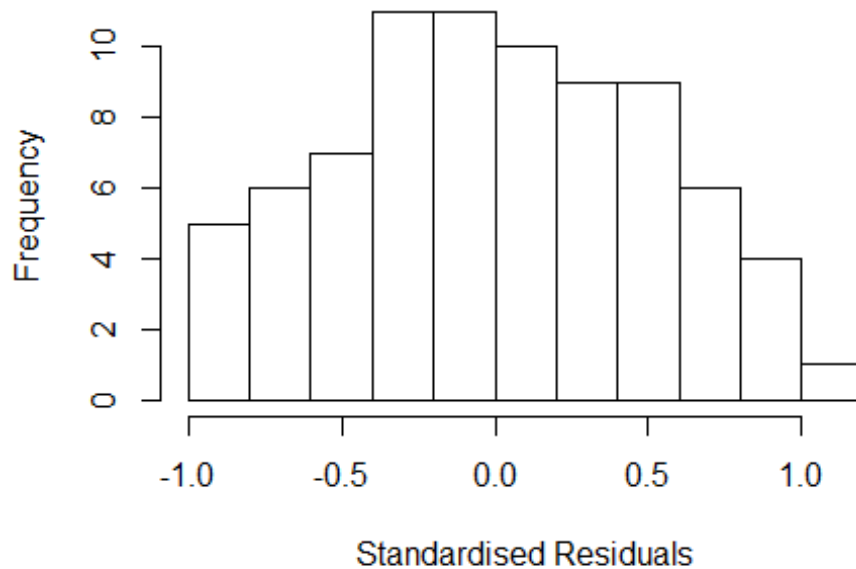
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9033 -0.2939 -0.0153 0.3786 1.0639
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6243 0.2511 2.49 0.0157 *
## FirstAuthorFemale1 0.0895 0.1975 0.45 0.6519
## Year1997 0.7857 0.8167 0.96 0.3398
## Year1998 0.1587 0.4509 0.35 0.7260
## Year1999 0.5162 0.4676 1.10 0.2739
## Year2000 0.6460 0.2979 2.17 0.0340 *
## Year2001 0.6734 0.5744 1.17 0.2456
## Year2002 0.2232 0.2559 0.87 0.3865
## Year2003 0.9989 0.3544 2.82 0.0065 **
## Year2004 0.5368 0.3237 1.66 0.1024
## Year2005 0.1956 0.3318 0.59 0.5577
## Year2006 0.7018 0.3509 2.00 0.0499 *
```

```

## Year2007          0.3463      0.2789      1.24      0.2191
## Year2008          0.6331      0.3079      2.06      0.0441 *
## Year2009          0.2795      0.2989      0.94      0.3534
## Year2010          0.5730      0.2980      1.92      0.0592 .
## Year2011          0.3813      0.3505      1.09      0.2810
## Year2012          0.6284      0.4146      1.52      0.1347
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.195, Adjusted R-squared:  -0.0299
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.737  0.905   0.960   0.935   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      1.27e-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 1.827 1      1.352
## Year            1.827 16      1.019

```


Residuals from last author



```
## [1] "List of 0 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
## -0.88096 -0.29580 -0.00179 0.39840 1.04500
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.624 0.251 2.49 0.0156 *
## LastAuthorFemale1 0.283 0.210 1.35 0.1830
## Year1997 0.786 0.814 0.96 0.3384
## Year1998 0.159 0.450 0.35 0.7254
## Year1999 0.516 0.466 1.11 0.2728
## Year2000 0.549 0.336 1.63 0.1075
## Year2001 0.674 0.574 1.17 0.2451
## Year2002 0.223 0.256 0.87 0.3863
## Year2003 1.018 0.345 2.95 0.0045 **
## Year2004 0.458 0.333 1.38 0.1734
## Year2005 0.196 0.332 0.59 0.5575
## Year2006 0.741 0.355 2.09 0.0412 *
```

```

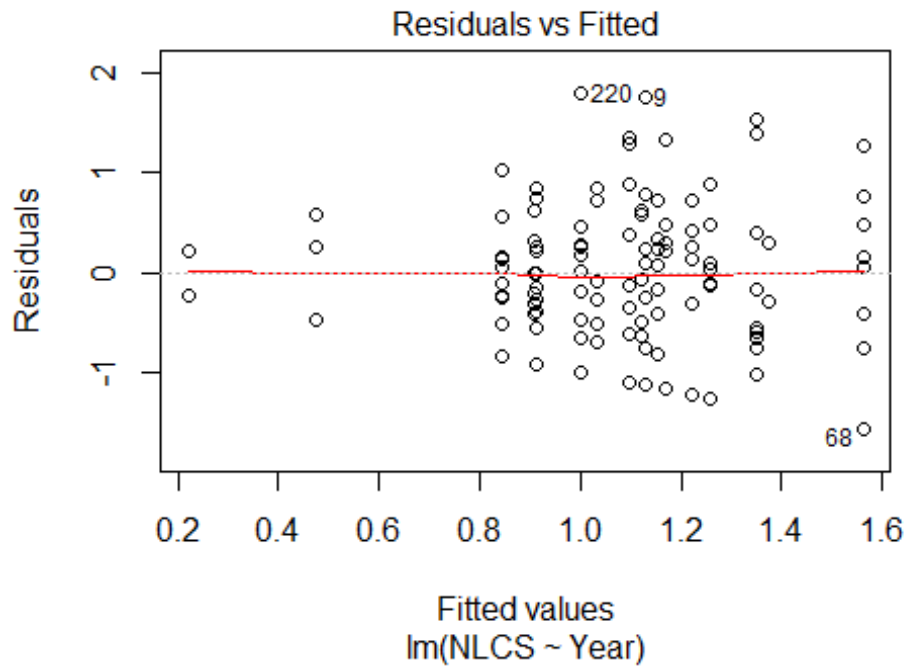
## Year2007          0.300      0.288      1.04      0.3023
## Year2008          0.650      0.328      1.98      0.0524 .
## Year2009          0.291      0.295      0.99      0.3272
## Year2010          0.551      0.298      1.85      0.0692 .
## Year2011          0.341      0.314      1.09      0.2811
## Year2012          0.628      0.414      1.52      0.1345
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.204, Adjusted R-squared:  -0.0183
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 74 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.746  0.910  0.959  0.937  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      1.27e-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: 79"
## [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
##    3    3    2    4    2    2    2    4    1    3    4    1    2    3    2
## 2011 2012
##    4    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    1    3    1    1    1    3    0    2    3    0    2    3    1
## 2011 2012

```

```

##      2      0
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      2      0      2      1      1      1      3      0      2      3      0      2      2      1
## 2011 2012
##      2      0
## [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: 2.6207413942089"
## [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: 23"
## [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
##      2      5     11     14     12     11     12     10     16      8     11     10     11     15     19
## 2011 2012
##     11     11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      2      8     11      5      8     10      7     10      6      6      7      6      6     11
## 2011 2012
##     10      6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      2      7      9      5      8      9      6     10      4      4      6      6      5      7
## 2011 2012
##      7      4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6

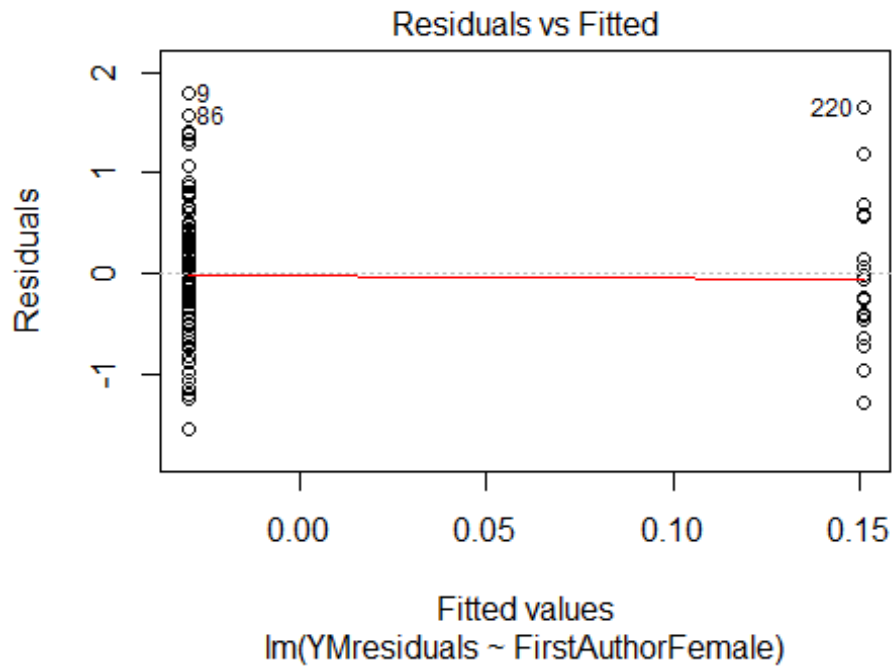
```



```
##
## 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: 4"
## [1] "Male first author team size 2018 geometric mean: 1.73205080756888"

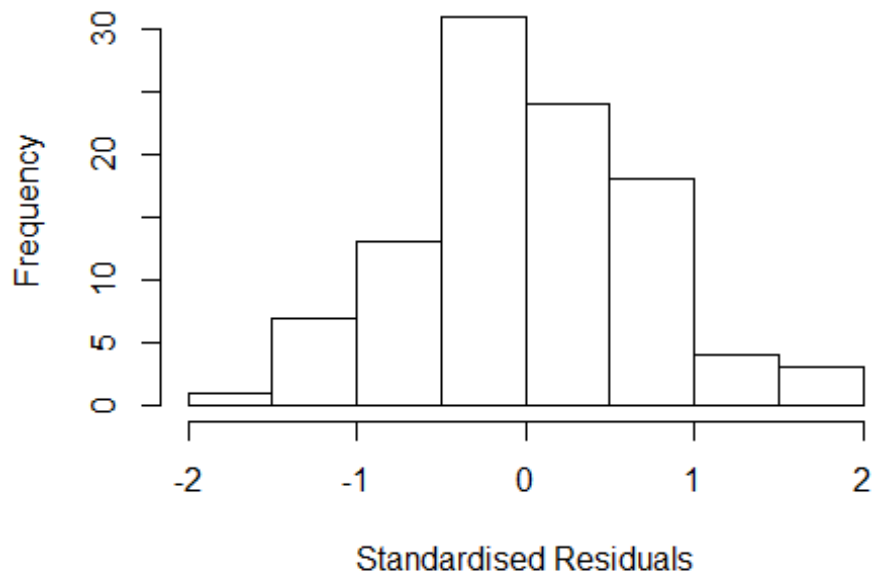
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 2.04767251107922"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF ^{1/(2*Df)}
FirstAuthorFemale	6.352	1	2.520
LastAuthorFemale	4.320	1	2.078
UniqueAuthors	30.564	4	1.533
Year	110.453	16	1.158

Residuals from first and last author and team size



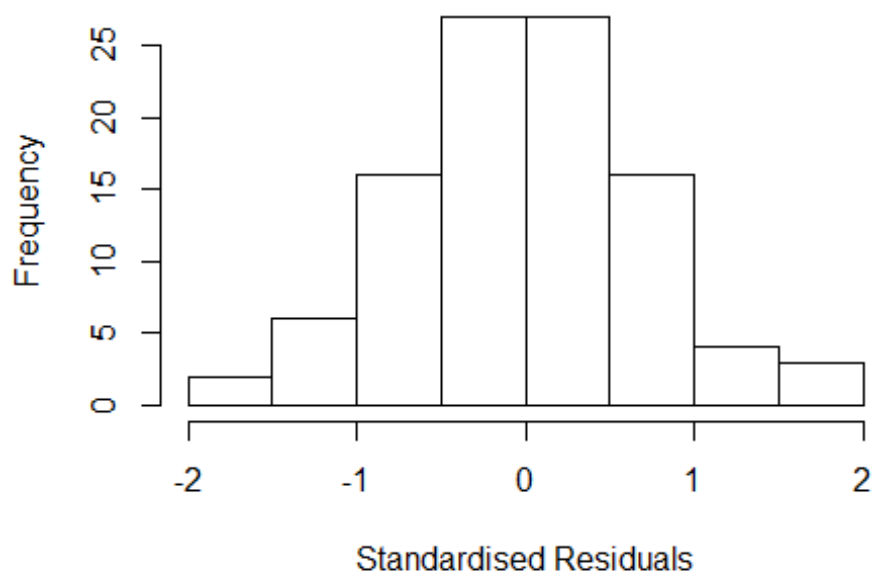
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6660 -0.4452 -0.0193 0.4614 1.8051
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.22000 0.15793 1.39 0.16756
## FirstAuthorFemale1 -0.00515 0.39896 -0.01 0.98974
## LastAuthorFemale1 0.09110 0.29907 0.30 0.76148
## UniqueAuthors2 0.03556 0.21901 0.16 0.87144
## UniqueAuthors3 -0.09292 0.24632 -0.38 0.70703
## UniqueAuthors4 0.12055 0.35080 0.34 0.73204
## UniqueAuthors5 -0.87231 0.53823 -1.62 0.10911
## Year1997 1.15300 0.26751 4.31 4.7e-05 ***
## Year1998 0.95480 0.38961 2.45 0.01650 *
## Year1999 0.79313 0.25143 3.15 0.00228 **
```

```

## Year2000      0.87217    0.32351    2.70  0.00859 **
## Year2001      1.41040    0.36611    3.85  0.00024 ***
## Year2002      1.06861    0.41293    2.59  0.01151 *
## Year2003      1.05373    0.39487    2.67  0.00926 **
## Year2004      0.82577    0.42789    1.93  0.05726 .
## Year2005      0.04875    0.35006    0.14  0.88960
## Year2006      0.80062    0.48304    1.66  0.10145
## Year2007      0.83201    0.30525    2.73  0.00792 **
## Year2008      0.65141    0.28487    2.29  0.02493 *
## Year2009      0.90127    0.40881    2.20  0.03043 *
## Year2010      1.16031    0.40226    2.88  0.00507 **
## Year2011      0.55059    0.25994    2.12  0.03735 *
## Year2012      1.28305    0.56481    2.27  0.02586 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.769
## Multiple R-squared:  0.174, Adjusted R-squared:  -0.0585
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.561  0.913  0.961  0.925  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      9.90e-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 7.455 1      2.730
## LastAuthorFemale 4.099 1      2.025
## Year      5.086 16      1.052

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.6089 -0.4519 -0.0157 0.4372 1.7858
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2200 0.1583 1.39 0.16829
## FirstAuthorFemale1 -0.0643 0.4598 -0.14 0.88915
## LastAuthorFemale1 0.1286 0.3228 0.40 0.69141
## Year1997 1.1530 0.2684 4.30 4.8e-05 ***
## Year1998 0.9477 0.3969 2.39 0.01924 *
## Year1999 0.8149 0.2213 3.68 0.00041 ***
## Year2000 0.8871 0.2866 3.10 0.00269 **
## Year2001 1.3889 0.3631 3.83 0.00025 ***
## Year2002 1.0400 0.4050 2.57 0.01205 *
## Year2003 1.0857 0.3291 3.30 0.00144 **
## Year2004 0.8188 0.3873 2.11 0.03755 *
## Year2005 0.0220 0.2819 0.08 0.93802
```

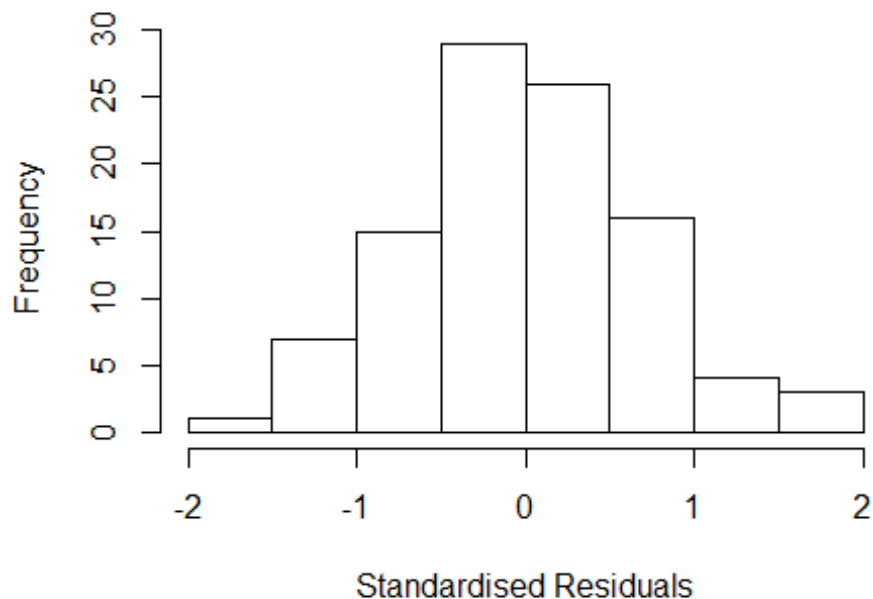


```

## Year2006      0.8261      0.3856      2.14  0.03511 *
## Year2007      0.8511      0.2642      3.22  0.00183 **
## Year2008      0.6518      0.2383      2.73  0.00765 **
## Year2009      0.8899      0.4346      2.05  0.04382 *
## Year2010      0.8565      0.3617      2.37  0.02024 *
## Year2011      0.5536      0.2478      2.23  0.02817 *
## Year2012      1.2868      0.6113      2.11  0.03833 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.72
## Multiple R-squared:  0.159, Adjusted R-squared:  -0.026
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 93 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.517  0.897  0.956  0.914  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      9.90e-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 3.164 1      1.779
## Year      3.164 16      1.037

```

Residuals from first author



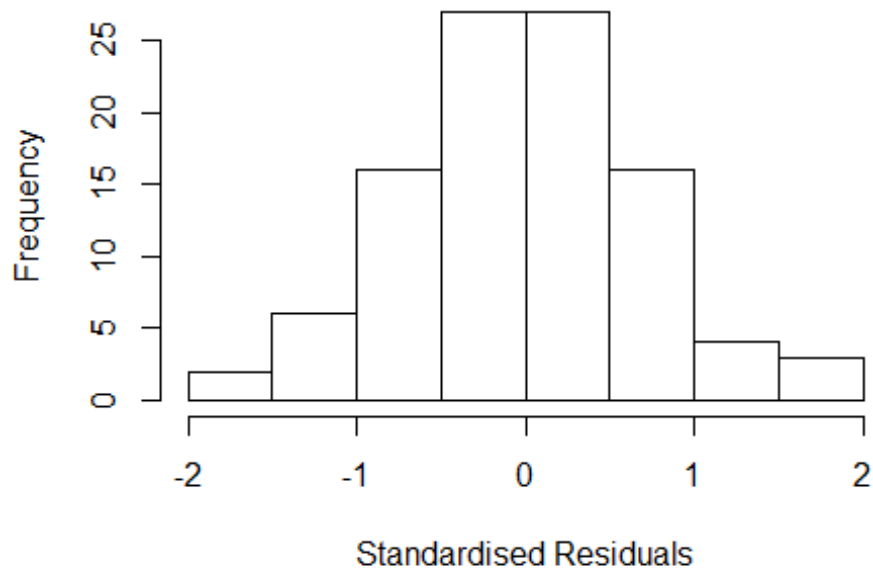
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6102 -0.4532 -0.0253 0.4392 1.7288
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2200 0.1583 1.39 0.16841
## FirstAuthorFemale1 0.0242 0.3105 0.08 0.93816
## Year1997 1.1530 0.2686 4.29 4.8e-05 ***
## Year1998 0.9382 0.4100 2.29 0.02465 *
## Year1999 0.8185 0.2207 3.71 0.00037 ***
## Year2000 0.8936 0.3090 2.89 0.00489 **
## Year2001 1.3902 0.3639 3.82 0.00026 ***
## Year2002 1.0364 0.4066 2.55 0.01264 *
## Year2003 1.0953 0.3293 3.33 0.00131 **
## Year2004 0.8168 0.3893 2.10 0.03895 *
## Year2005 0.0215 0.2820 0.08 0.93952
## Year2006 0.8271 0.3863 2.14 0.03518 *
```

```

## Year2007          0.8844      0.2591      3.41  0.00099 ***
## Year2008          0.6792      0.2283      2.98  0.00383 **
## Year2009          0.8925      0.4414      2.02  0.04639 *
## Year2010          0.8591      0.3526      2.44  0.01696 *
## Year2011          0.5661      0.2501      2.26  0.02624 *
## Year2012          1.2635      0.5902      2.14  0.03524 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.711
## Multiple R-squared:  0.157, Adjusted R-squared:  -0.0152
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 93 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.534  0.892  0.956  0.912  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      9.90e-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 1.551 1      1.246
## Year      1.551 16      1.014

```

Residuals from last author



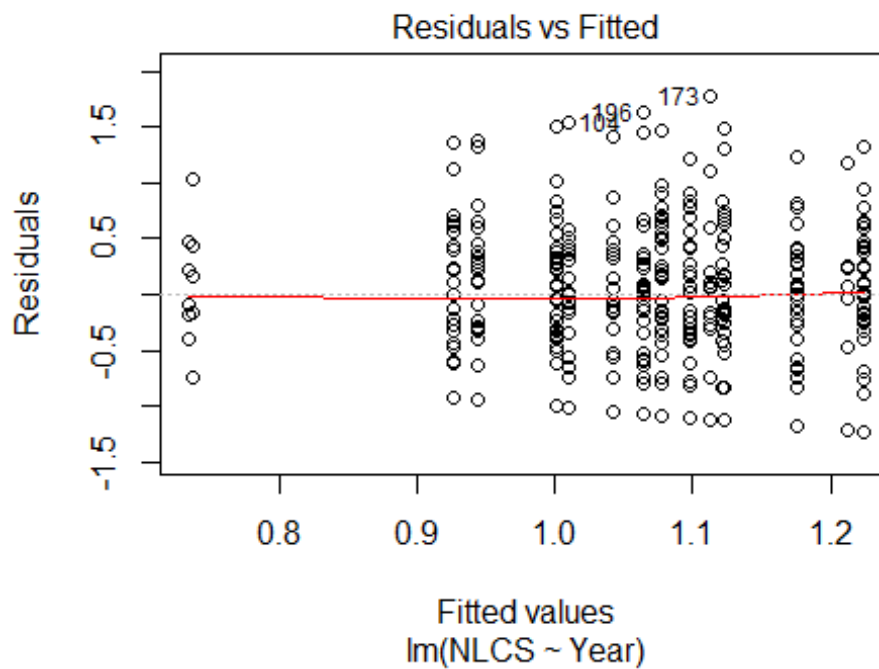
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6090 -0.4623 -0.0107 0.4374 1.7250
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.2200 0.1583 1.39 0.16826
## LastAuthorFemale1 0.0881 0.2008 0.44 0.66183
## Year1997 1.1530 0.2684 4.30 4.7e-05 ***
## Year1998 0.9420 0.4063 2.32 0.02290 *
## Year1999 0.8128 0.2223 3.66 0.00045 ***
## Year2000 0.8819 0.2905 3.04 0.00321 **
## Year2001 1.3890 0.3636 3.82 0.00026 ***
## Year2002 1.0397 0.4089 2.54 0.01287 *
## Year2003 1.0807 0.3249 3.33 0.00131 **
## Year2004 0.8186 0.3891 2.10 0.03841 *
## Year2005 0.0219 0.2822 0.08 0.93821
## Year2006 0.8262 0.3862 2.14 0.03534 *
```

```

## Year2007          0.8516      0.2626      3.24  0.00170 **
## Year2008          0.6542      0.2381      2.75  0.00735 **
## Year2009          0.8604      0.3385      2.54  0.01287 *
## Year2010          0.8597      0.3489      2.46  0.01579 *
## Year2011          0.5463      0.2382      2.29  0.02437 *
## Year2012          1.2568      0.5135      2.45  0.01649 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.719
## Multiple R-squared:  0.158, Adjusted R-squared:  -0.0146
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 93 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.544  0.898  0.957  0.914  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      9.90e-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: 101"
## [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 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##   20   11   18   18   33   26   32   31   41   45   43   56   62   58   62
## 2012
##   62
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    6    5    7    9   20   14   26   20   28   24   23   31   36   36   33
## 2012

```

```
## 38
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 5 4 6 9 18 11 20 18 24 20 17 28 27 29 28
## 2012
## 32
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 7.3, df = 15, p-value = 0.9
```



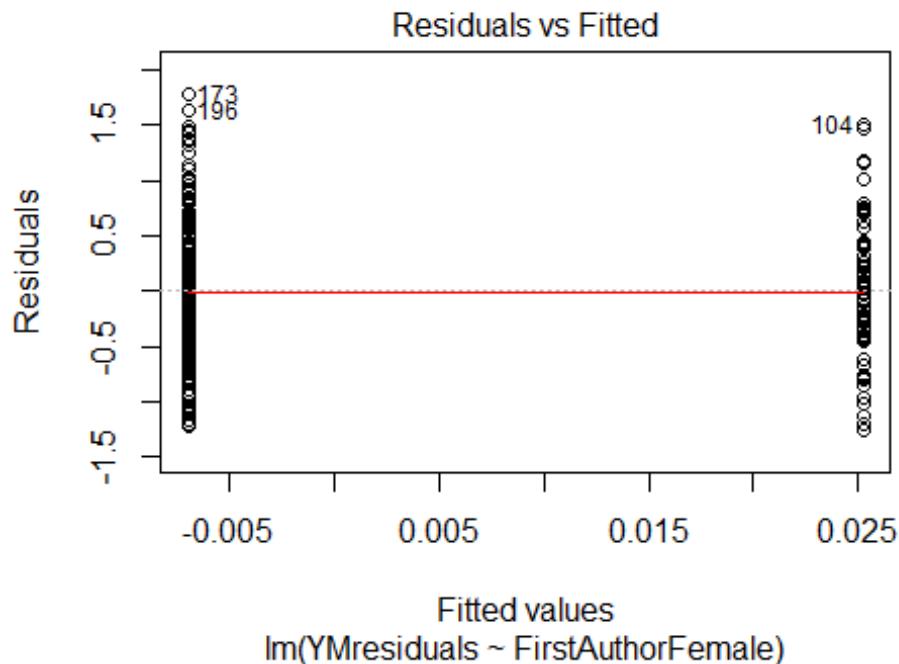
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.16, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 2.26574964973706"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

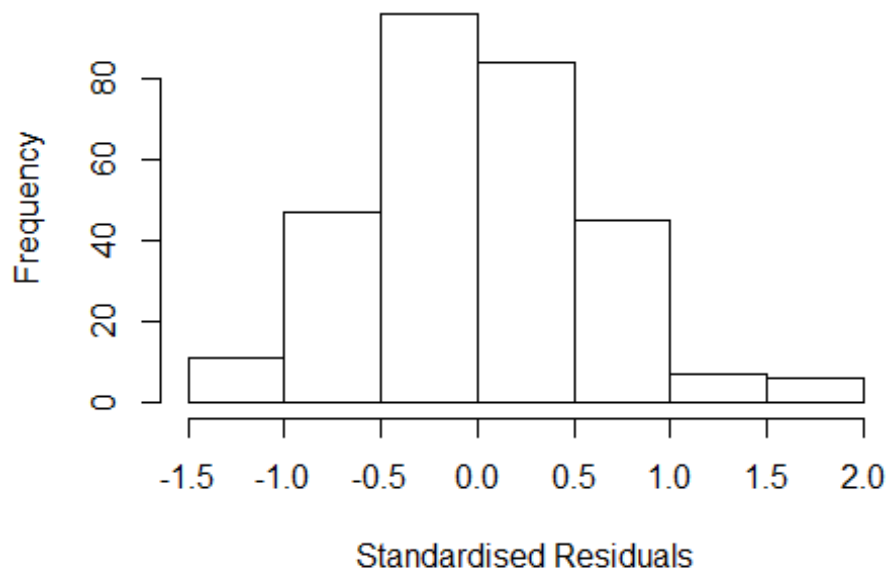
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 2.26574964973706"

## 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 = 11, 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 1.330  1      1.153
## LastAuthorFemale  1.280  1      1.131
## UniqueAuthors    2.572  4      1.125
## Year              3.529 15      1.043
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.1913 -0.3763 -0.0257 0.4111 1.8508
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8922 0.2779 3.21 0.0015 **
## FirstAuthorFemale1 0.1325 0.0902 1.47 0.1428
## LastAuthorFemale1 -0.1503 0.0928 -1.62 0.1064
## UniqueAuthors2 -0.0624 0.0942 -0.66 0.5083
## UniqueAuthors3 0.0936 0.1062 0.88 0.3786
## UniqueAuthors4 0.0274 0.1617 0.17 0.8656
## UniqueAuthors5 -0.0373 0.1404 -0.27 0.7906
## Year1998 -0.2509 0.2933 -0.86 0.3931
## Year1999 0.4665 0.3329 1.40 0.1622
## Year2000 0.2765 0.3330 0.83 0.4071
```

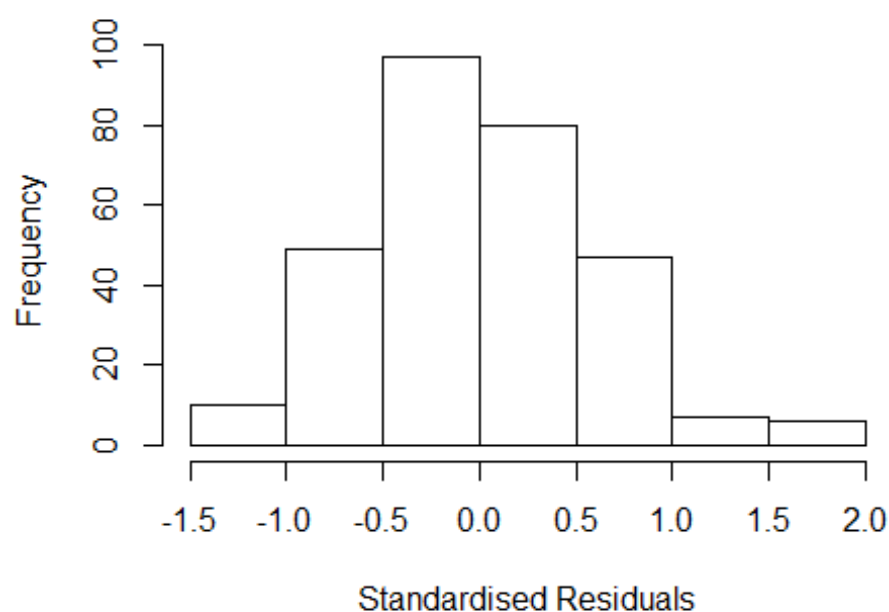


```

## Year2001          0.1349      0.3078      0.44      0.6617
## Year2002          0.0484      0.3798      0.13      0.8987
## Year2003          0.0562      0.3152      0.18      0.8586
## Year2004          0.0657      0.3142      0.21      0.8346
## Year2005          0.0224      0.3105      0.07      0.9426
## Year2006         -0.0402      0.3230     -0.12      0.9011
## Year2007          0.2383      0.3245      0.73      0.4634
## Year2008          0.2410      0.2974      0.81      0.4186
## Year2009          0.3558      0.3057      1.16      0.2456
## Year2010          0.2720      0.3119      0.87      0.3840
## Year2011          0.1882      0.3035      0.62      0.5357
## Year2012         -0.0111      0.2990     -0.04      0.9705
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.0638, Adjusted R-squared:  -0.00795
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 272 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.384  0.876  0.956  0.917  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      3.38e-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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.269 1      1.127
## LastAuthorFemale  1.241 1      1.114
## Year              1.428 15      1.012

```

Residuals from first and last author



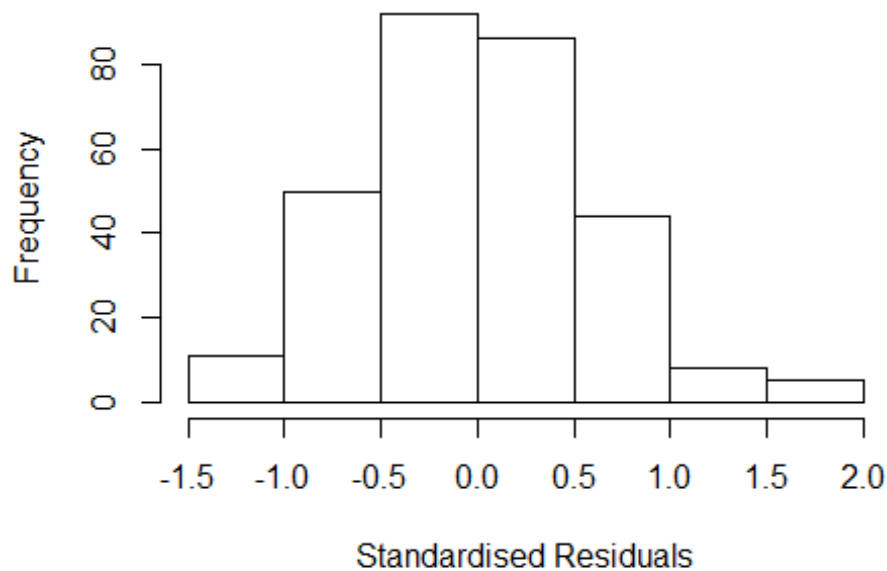
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1617 -0.3743 -0.0331 0.4146 1.9759
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.87058 0.28149 3.09 0.0022 **
## FirstAuthorFemale1 0.13249 0.09020 1.47 0.1430
## LastAuthorFemale1 -0.14174 0.09299 -1.52 0.1286
## Year1998 -0.22484 0.30360 -0.74 0.4596
## Year1999 0.48452 0.32625 1.49 0.1386
## Year2000 0.25194 0.32202 0.78 0.4347
## Year2001 0.16722 0.30736 0.54 0.5868
## Year2002 0.03852 0.36927 0.10 0.9170
## Year2003 0.09172 0.31194 0.29 0.7689
## Year2004 0.08550 0.30962 0.28 0.7826
## Year2005 0.03418 0.30629 0.11 0.9112
## Year2006 -0.01799 0.31572 -0.06 0.9546
```

```

## Year2007          0.24553    0.32249    0.76    0.4471
## Year2008          0.24195    0.29145    0.83    0.4072
## Year2009          0.38322    0.29932    1.28    0.2015
## Year2010          0.29114    0.30782    0.95    0.3451
## Year2011          0.19403    0.29347    0.66    0.5090
## Year2012          0.00649    0.29545    0.02    0.9825
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.629
## Multiple R-squared:  0.0575, Adjusted R-squared:  -8.46e-05
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 269 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.303  0.868  0.955  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      3.38e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.195 1      1.093
## Year              1.195 15      1.006

```

Residuals from first author

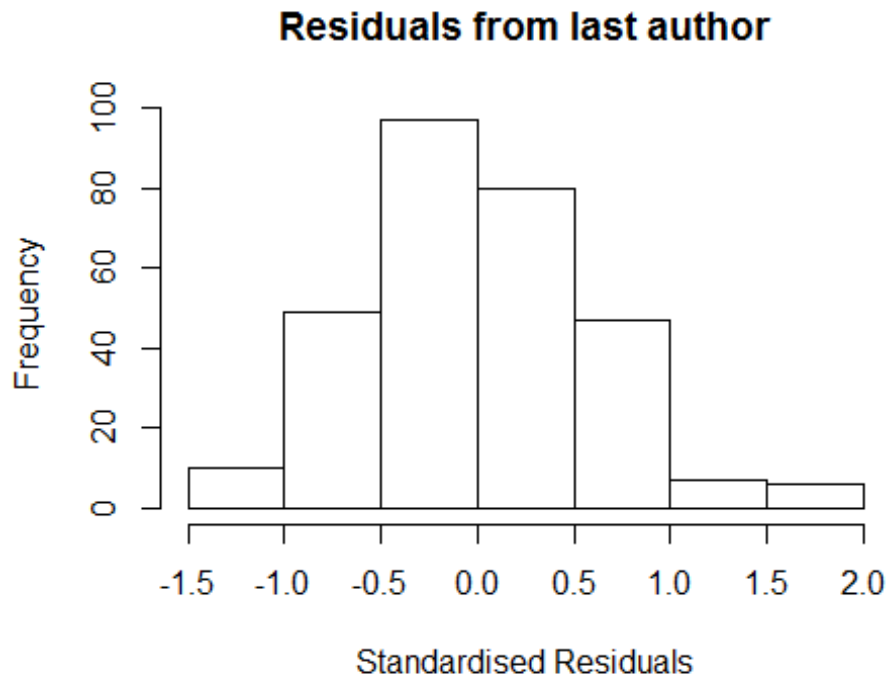


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2200 -0.3591 -0.0168 0.4240 1.9653
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.83250 0.27526 3.02 0.0027 **
## FirstAuthorFemale1 0.08338 0.08861 0.94 0.3475
## Year1998 -0.19985 0.30901 -0.65 0.5183
## Year1999 0.50082 0.32919 1.52 0.1293
## Year2000 0.27807 0.31556 0.88 0.3790
## Year2001 0.18792 0.30510 0.62 0.5384
## Year2002 0.08723 0.36283 0.24 0.8102
## Year2003 0.13432 0.30589 0.44 0.6609
## Year2004 0.11164 0.30461 0.37 0.7143
## Year2005 0.05110 0.30666 0.17 0.8678
## Year2006 0.00807 0.31011 0.03 0.9792
## Year2007 0.27114 0.31810 0.85 0.3947
```

```

## Year2008          0.26306    0.28971    0.91    0.3647
## Year2009          0.38751    0.29752    1.30    0.1938
## Year2010          0.30677    0.30619    1.00    0.3173
## Year2011          0.18532    0.29291    0.63    0.5275
## Year2012          0.02082    0.29240    0.07    0.9433
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.626
## Multiple R-squared:  0.0507, Adjusted R-squared:  -0.00378
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 268 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.303  0.865  0.955  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      3.38e-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"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.202  1         1.096
## Year             1.202 15         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
## -1.1686 -0.3722 -0.0249 0.4155 1.9422
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.93310 0.30796 3.03 0.0027 **
## LastAuthorFemale1 -0.09663 0.09309 -1.04 0.3002
## Year1998 -0.23420 0.34159 -0.69 0.4935
## Year1999 0.49322 0.36055 1.37 0.1724
## Year2000 0.20004 0.34334 0.58 0.5606
## Year2001 0.11782 0.33050 0.36 0.7217
## Year2002 0.00973 0.39289 0.02 0.9803
## Year2003 0.05522 0.34071 0.16 0.8714
## Year2004 0.03628 0.33585 0.11 0.9141
## Year2005 -0.01092 0.33202 -0.03 0.9738
## Year2006 -0.06984 0.34145 -0.20 0.8381
## Year2007 0.20858 0.34765 0.60 0.5490
```

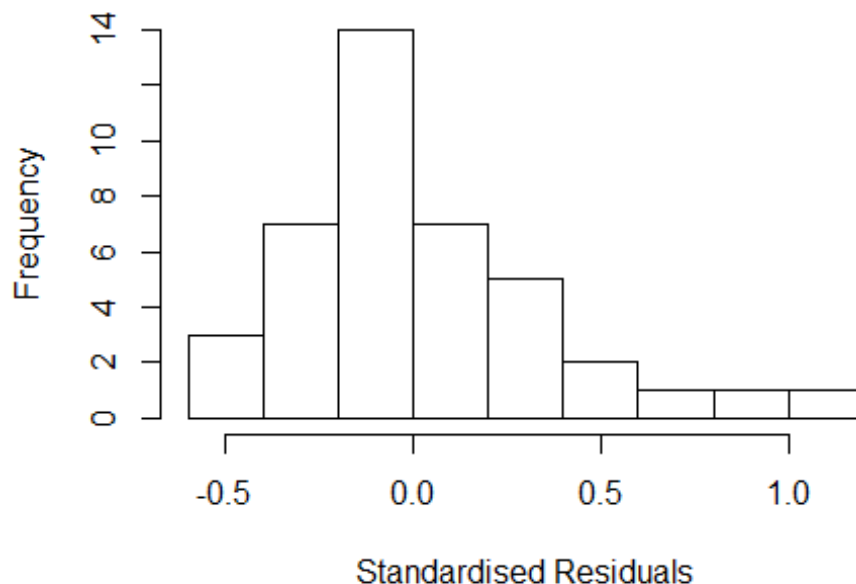
```

## Year2008          0.21915      0.32083      0.68      0.4951
## Year2009          0.32679      0.32453      1.01      0.3148
## Year2010          0.23554      0.33343      0.71      0.4805
## Year2011          0.14959      0.31823      0.47      0.6387
## Year2012         -0.04733      0.31992     -0.15      0.8825
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.631
## Multiple R-squared:  0.0517, Adjusted R-squared:  -0.00273
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 275 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.324  0.867  0.956  0.914  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      3.38e-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: 296"
## [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
##    3    5    7    6    3   10   11    4    4    3    4    8    7    3    8
## 2011 2012
##    7    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    3    5    3    1    5    6    2    2    2    2    3    2    2    6
## 2011 2012
##    3    4

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    3    3    3    0    4    6    1    1    2    1    2    2    1    5
## 2011 2012
##    3    4
## [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: 2.37841423000544"
## [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"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.636e+13 1      4044183
## LastAuthorFemale -6.271e+14 1          NaN
## Year              -8.633e-02 14          NaN
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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"
```

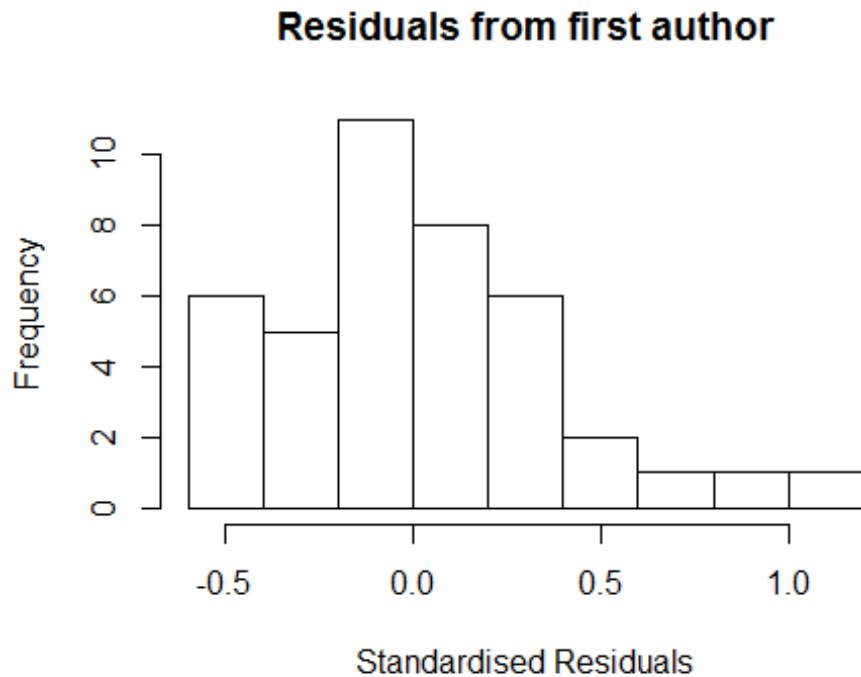


```

## Residuals:
##      Min       1Q   Median       3Q      Max
## -5.09e-01 -1.74e-01  5.55e-17  1.29e-01  1.13e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.08959    0.43318    2.52   0.019 *
## FirstAuthorFemale1  0.28444    0.21475    1.32   0.198
## LastAuthorFemale1 -1.42144    0.21475   -6.62  7.6e-07 ***
## Year1998         0.12252    0.43493    0.28   0.781
## Year1999         0.27287    0.50423    0.54   0.593
## Year2001        -0.02493    0.45535   -0.05   0.957
## Year2002        -0.00732    0.45854   -0.02   0.987
## Year2003        -0.14059    0.43318   -0.32   0.748
## Year2004        -1.08959    0.43318   -2.52   0.019 *
## Year2005         0.15191    0.46177    0.33   0.745
## Year2006        -0.35859    0.43318   -0.83   0.416
## Year2007        -0.20359    0.43331   -0.47   0.643
## Year2008         0.36141    0.43318    0.83   0.412
## Year2009        -0.11359    0.43318   -0.26   0.795
## Year2010        -0.61307    0.45920   -1.34   0.194
## Year2011        -0.03291    0.46524   -0.07   0.944
## Year2012        -0.58108    0.48802   -1.19   0.245
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.55, Adjusted R-squared:  0.251
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~ = 1. The remaining 31 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.377  0.903  0.971   0.917   0.986   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      2.44e-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 7.21  1      2.685
## Year              7.21 14      1.073
```



```
## [1] "List of 0 outliers 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
## -5.29e-01 -2.13e-01 -1.11e-16  2.12e-01  1.11e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.1094     0.4234   2.62    0.015 *
## FirstAuthorFemale1 -0.0782     0.3398  -0.23    0.820
## Year1998         0.1027     0.4252   0.24    0.811
## Year1999         0.4019     0.5053   0.80    0.434
## Year2001         0.0459     0.4581   0.10    0.921
## Year2002        -0.0278     0.4461  -0.06    0.951
## Year2003        -0.1604     0.4234  -0.38    0.708
```

```

## Year2004          -1.1094      0.4234   -2.62    0.015 *
## Year2005           0.1321      0.4523    0.29    0.773
## Year2006          -0.3784      0.4234   -0.89    0.380
## Year2007          -0.2234      0.4236   -0.53    0.602
## Year2008          -0.1879      0.7318   -0.26    0.800
## Year2009          -0.1334      0.4234   -0.32    0.755
## Year2010          -0.6156      0.4571   -1.35    0.190
## Year2011          -0.0505      0.4589   -0.11    0.913
## Year2012          -0.5983      0.4782   -1.25    0.223
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.455, Adjusted R-squared:  0.128
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 33 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.445  0.882  0.969   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      2.44e-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 -0.009719 1      NaN
## Year             -0.009719 14      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
## -5.11e-01 -1.89e-01  5.00e-16  2.10e-01  1.11e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.11e+00   4.23e-01   2.62e+00   0.015 *
## LastAuthorFemale1 -1.14e+00   1.27e-15 -8.99e+14  <2e-16 ***
## Year1998         1.04e-01   4.25e-01   2.50e-01   0.808
## Year1999         3.72e-01   4.86e-01   7.70e-01   0.451
## Year2001         2.92e-02   4.63e-01   6.00e-02   0.950
## Year2002        -2.63e-02   4.46e-01  -6.00e-02   0.953
## Year2003        -1.59e-01   4.23e-01  -3.80e-01   0.710
## Year2004        -1.11e+00   4.23e-01  -2.62e+00   0.015 *
## Year2005         1.34e-01   4.52e-01   3.00e-01   0.770
## Year2006        -3.77e-01   4.23e-01  -8.90e-01   0.381
## Year2007        -2.22e-01   4.23e-01  -5.20e-01   0.605
## Year2008         3.43e-01   4.23e-01   8.10e-01   0.425
## Year2009        -1.32e-01   4.23e-01  -3.10e-01   0.758
## Year2010        -6.15e-01   4.57e-01  -1.35e+00   0.190
## Year2011        -4.92e-02   4.59e-01  -1.10e-01   0.915
## Year2012        -5.97e-01   4.78e-01  -1.25e+00   0.224
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.408
## Multiple R-squared:  0.528, Adjusted R-squared:  0.244
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 30 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.439  0.897  0.969  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      2.44e-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: 41"

```

```

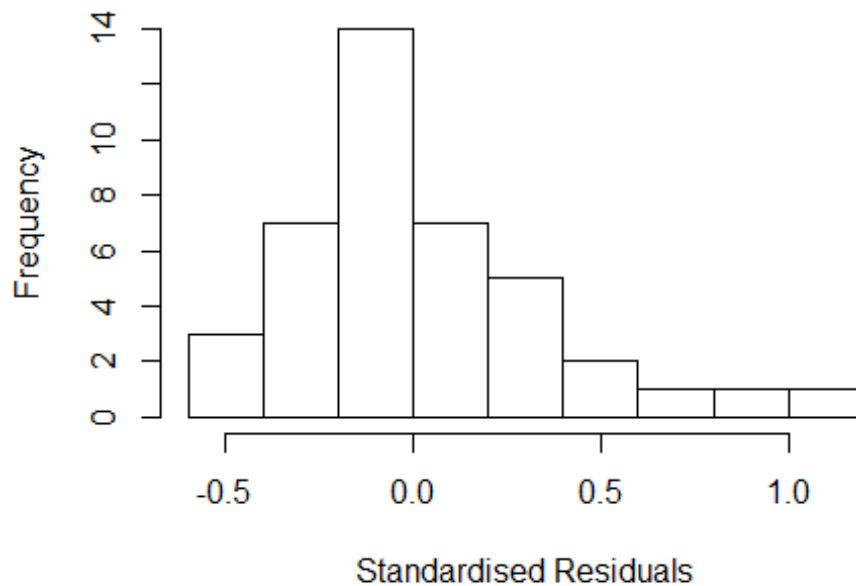
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1708"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    3    3    4    8    5    1    3    4    8    6    7    6    11
## 2011 2012
##    7    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    2    2    3    3    3    1    2    3    4    5    3    3    7
## 2011 2012
##    5    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    1    3    3    3    1    1    2    2    3    2    3    5
## 2011 2012
##    4    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.46410161513775"
## [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"

## 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

```

Residuals from last author



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1 NaN
## LastAuthorFemale NaN 1 NaN
## Year NaN 15 NaN

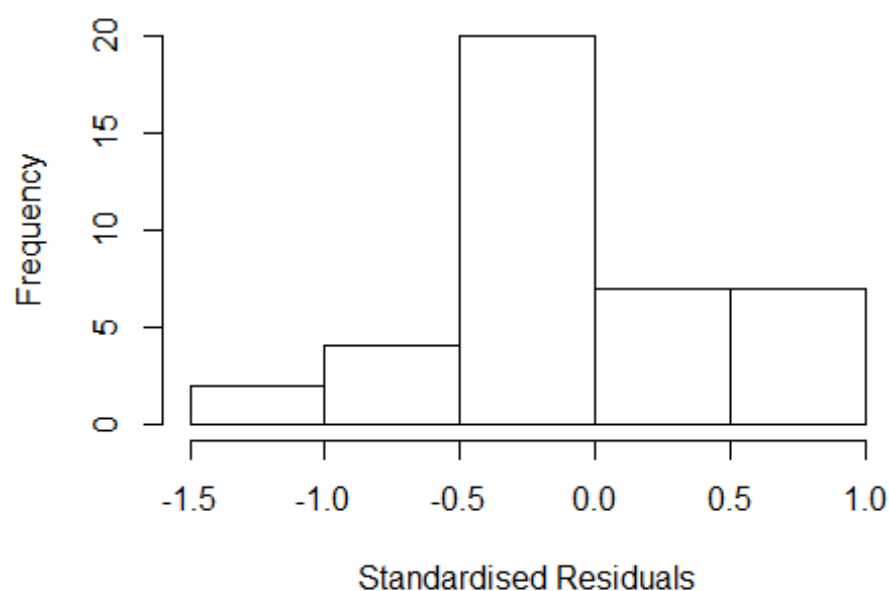
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3927 -0.1647 -0.0472  0.1894  0.9575
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0760     0.0000      Inf <2e-16 ***
## FirstAuthorFemale1  0.2154     0.3161    0.68  0.503
## LastAuthorFemale1 -0.2135     0.4285   -0.50  0.623
## Year1998          -0.7230     0.0000   -Inf <2e-16 ***
## Year1999          -0.3200     0.0000   -Inf <2e-16 ***
## Year2000           0.1601     0.3484    0.46  0.650
## Year2001           0.8437     0.7305    1.15  0.261
```

```

## Year2002          -0.2278      0.3362    -0.68      0.505
## Year2003           0.8540      0.0000      Inf    <2e-16 ***
## Year2004          -0.1160      0.0000    -Inf    <2e-16 ***
## Year2005           0.3308      0.1666      1.99      0.060 .
## Year2006          -0.3375      0.1311     -2.57      0.017 *
## Year2007          -0.2320      0.3061     -0.76      0.457
## Year2008          -0.2287      0.2422     -0.94      0.355
## Year2009           0.3171      0.3880      0.82      0.422
## Year2010          -0.0705      0.3358     -0.21      0.836
## Year2011          -0.0871      0.1254     -0.69      0.495
## Year2012           0.2419      0.3167      0.76      0.453
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.619
## Multiple R-squared:  0.373, Adjusted R-squared:  -0.111
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.592  0.896  0.976  0.936  0.994  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.50e-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"
## 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

```

Residuals from first and last author



```
##              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
## -1.3702 -0.1858 -0.0165  0.2407  0.9750
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.08e+00   0.00e+00      Inf <2e-16 ***
## FirstAuthorFemale1 1.66e-01   3.28e-01   5.10e-01   0.617
## Year1998        -7.23e-01   0.00e+00    -Inf <2e-16 ***
## Year1999        -3.20e-01   2.94e-08  -1.09e+07 <2e-16 ***
## Year2000         9.68e-02   3.02e-01   3.20e-01   0.751
## Year2001         8.21e-01   7.13e-01   1.15e+00   0.261
## Year2002        -2.26e-01   3.33e-01  -6.80e-01   0.504
## Year2003         8.54e-01   1.05e-08   8.16e+07 <2e-16 ***
```

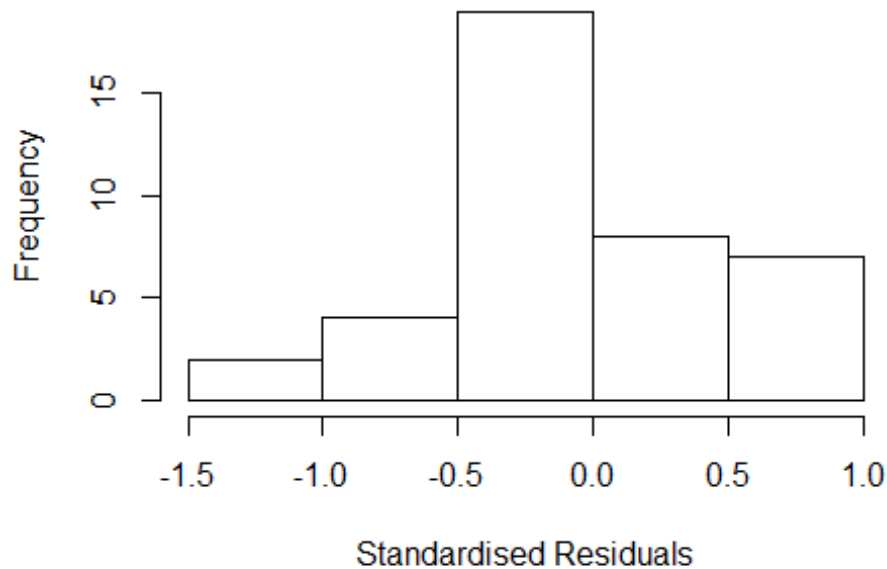


```

## Year2004          -1.16e-01    0.00e+00      -Inf    <2e-16 ***
## Year2005          3.55e-01    1.68e-01    2.12e+00    0.045 *
## Year2006          -3.38e-01    1.31e-01   -2.58e+00    0.017 *
## Year2007          -2.83e-01    3.30e-01   -8.60e-01    0.400
## Year2008          -3.36e-01    1.93e-01   -1.74e+00    0.095 .
## Year2009          3.49e-01    4.01e-01    8.70e-01    0.394
## Year2010          -7.32e-02    3.30e-01   -2.20e-01    0.827
## Year2011          -8.67e-02    1.25e-01   -6.90e-01    0.496
## Year2012          2.83e-01    3.34e-01    8.50e-01    0.405
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.648
## Multiple R-squared:  0.362, Adjusted R-squared:  -0.0825
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~ = 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.634 0.908 0.959 0.933 0.993 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      2.50e-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

```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 2.976 1          1.725
## Year            2.976 15          1.037

## [1] "List of 0 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
## -1.4330 -0.1859 -0.0135  0.1859  1.0727
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.08e+00  0.00e+00      Inf < 2e-16 ***
## LastAuthorFemale1 -1.37e-01  3.75e-01 -3.70e-01  0.718
## Year1998       -7.23e-01  2.72e-08 -2.65e+07 < 2e-16 ***
## Year1999       -3.20e-01  2.74e-08 -1.17e+07 < 2e-16 ***
## Year2000        1.38e-01  3.34e-01  4.10e-01  0.683
## Year2001        8.84e-01  7.53e-01  1.17e+00  0.252
## Year2002       -2.31e-01  3.41e-01 -6.80e-01  0.506
## Year2003        8.54e-01  2.66e-08  3.21e+07 < 2e-16 ***
```

```

## Year2004          -1.16e-01    5.68e-09 -2.04e+07 < 2e-16 ***
## Year2005          4.38e-01    2.37e-02  1.85e+01 2.6e-15 ***
## Year2006          -3.38e-01    1.31e-01 -2.57e+00 0.017 *
## Year2007          -1.88e-01    2.81e-01 -6.70e-01 0.511
## Year2008          -2.67e-01    2.35e-01 -1.14e+00 0.267
## Year2009          4.17e-01    5.25e-01  7.90e-01 0.435
## Year2010          -6.61e-02    3.44e-01 -1.90e-01 0.849
## Year2011          -8.77e-02    1.26e-01 -7.00e-01 0.493
## Year2012          4.12e-01    2.21e-01  1.87e+00 0.075 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.581
## Multiple R-squared:  0.376, Adjusted R-squared:  -0.0578
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 30 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.523 0.888 0.953 0.917 0.992 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           2.50e-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: 40"
## [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
##    2    3    6    2    4    6    5    3    7    4    3   11    9    9    5
## 2011 2012
##    7   13
##

```

```

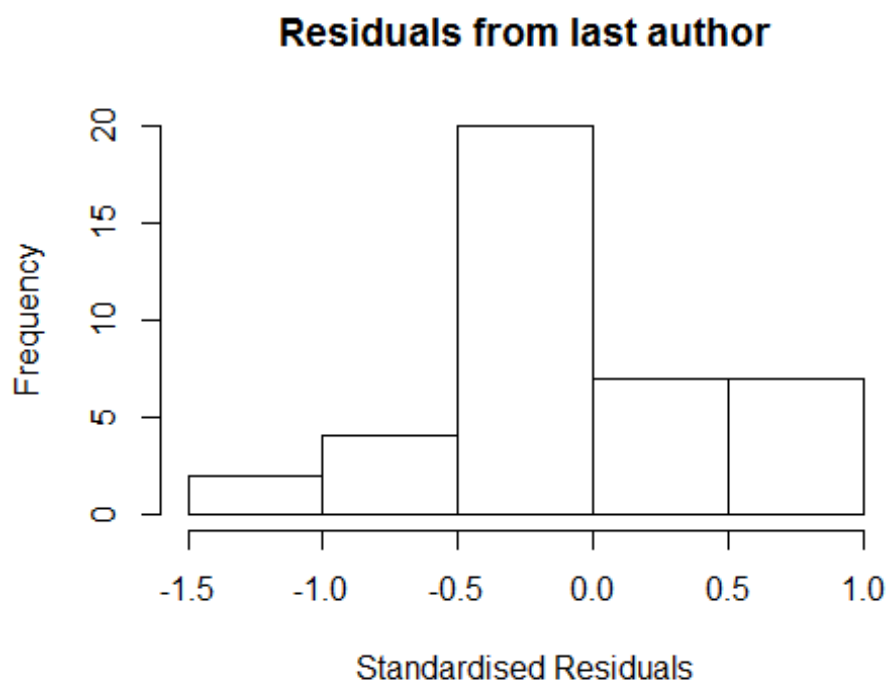
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      4      2      2      4      4      3      3      2      2      5      4      7      3
## 2011 2012
##      4      8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      3      2      2      4      4      3      2      2      1      4      3      7      2
## 2011 2012
##      1      6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 2.51984209978975"

## 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 = 2.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 2.82842712474619"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

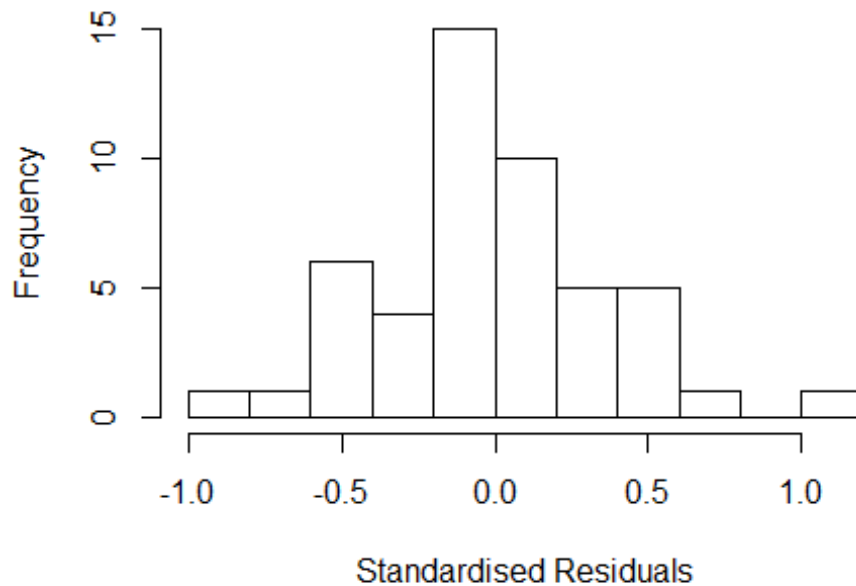
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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	3.737e+14	1	1.933e+07
## LastAuthorFemale	1.195e+15	1	3.457e+07
## UniqueAuthors	2.909e+43	4	2.710e+05
## Year	2.873e+46	16	2.830e+01

Residuals from first and last author and team size



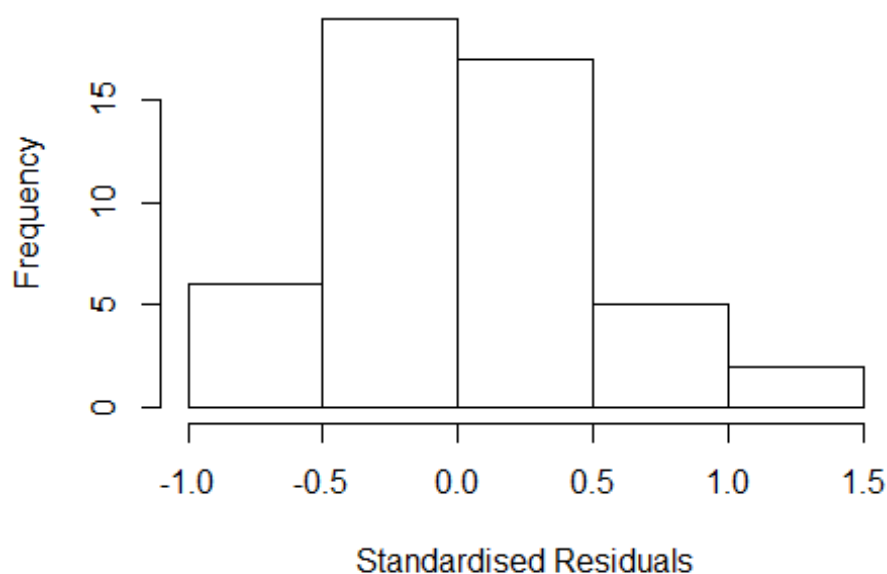
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.24e-01 -1.85e-01 -1.11e-16  1.85e-01  1.03e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.5835     0.0280   56.62 < 2e-16 ***
## FirstAuthorFemale1  0.3105     0.2587    1.20  0.2409
## LastAuthorFemale1 -0.6916     0.5953   -1.16  0.2559
## UniqueAuthors2     0.1971     0.2104    0.94  0.3576
## UniqueAuthors3    -0.4096     0.4811   -0.85  0.4023
## UniqueAuthors4     1.1363     0.5838    1.95  0.0625 .
## UniqueAuthors5     0.0872     0.2428    0.36  0.7224
## Year1997          -0.3756     0.2122   -1.77  0.0886 .
## Year1998          -0.8173     0.2752   -2.97  0.0063 **
## Year1999          -0.7660     0.2197   -3.49  0.0018 **
```

```

## Year2000          -1.2753      0.2380      -5.36      1.3e-05 ***
## Year2001           0.4593      0.2444       1.88      0.0714 .
## Year2002          -0.1850      0.3268      -0.57      0.5762
## Year2003          -0.7599      0.5670      -1.34      0.1918
## Year2004          -0.6792      0.4393      -1.55      0.1342
## Year2005          -0.4113      0.4405      -0.93      0.3590
## Year2006           0.5266      0.4191       1.26      0.2201
## Year2007          -0.3894      0.2392      -1.63      0.1156
## Year2008          -0.4917      0.4635      -1.06      0.2985
## Year2009          -0.4643      0.7768      -0.60      0.5552
## Year2010          -0.6175      0.4224      -1.46      0.1558
## Year2011          -0.1297      0.2444      -0.53      0.6001
## Year2012          -0.8058      0.1665      -4.84      5.1e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.455
## Multiple R-squared:  0.638, Adjusted R-squared:  0.331
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.591  0.908  0.967  0.932  0.993  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.04e-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 6.383e+14 1      2.527e+07
## LastAuthorFemale  2.452e+15 1      4.951e+07
## Year              1.525e+16 16      3.204e+00

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -9.50e-01 -2.26e-01 3.33e-16 2.19e-01 1.35e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5835 0.0280 56.62 < 2e-16 ***
## FirstAuthorFemale1 0.2516 0.2395 1.05 0.30189
## LastAuthorFemale1 -0.4910 0.4694 -1.05 0.30392
## Year1997 -0.1785 0.0280 -6.38 4.8e-07 ***
## Year1998 -0.7040 0.2700 -2.61 0.01407 *
## Year1999 -0.8368 0.2074 -4.03 0.00035 ***
## Year2000 -1.1473 0.1294 -8.87 7.0e-10 ***
## Year2001 0.4526 0.2689 1.68 0.10277
## Year2002 -0.2313 0.2701 -0.86 0.39861
## Year2003 -0.6332 0.5276 -1.20 0.23951
## Year2004 -0.8840 0.1684 -5.25 1.2e-05 ***
## Year2005 -0.8918 0.2103 -4.24 0.00020 ***
```

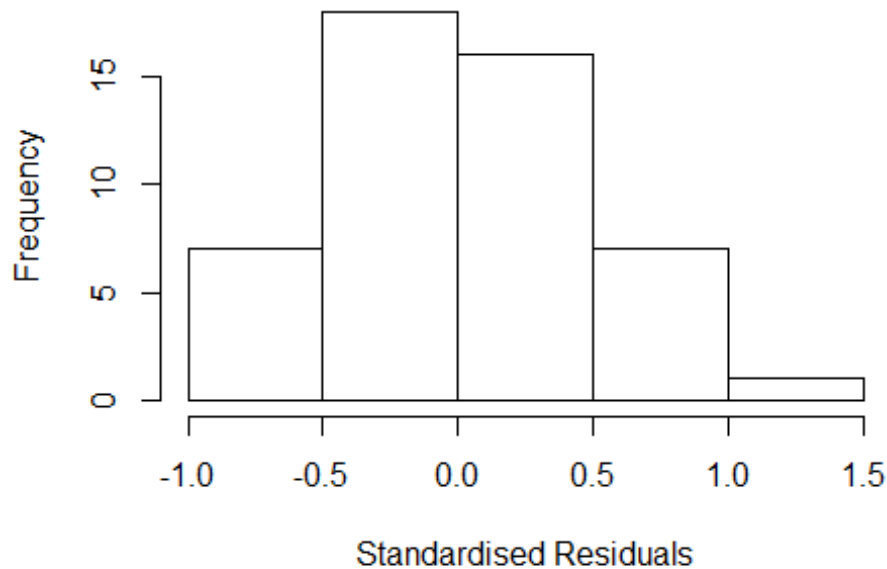


```

## Year2006          0.3849      0.3439      1.12  0.27191
## Year2007          -0.3194      0.2190     -1.46  0.15511
## Year2008          -0.5863      0.3847     -1.52  0.13804
## Year2009          -0.6181      0.4064     -1.52  0.13872
## Year2010          -0.5190      0.5385     -0.96  0.34282
## Year2011          -0.0425      0.0280     -1.52  0.13904
## Year2012          -0.7002      0.1309     -5.35  8.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.496, Adjusted R-squared:  0.194
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.417  0.898  0.968  0.917  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      2.04e-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 -7.739e+15  1      NaN
## Year              -7.739e+15 16      NaN

```

Residuals from first author



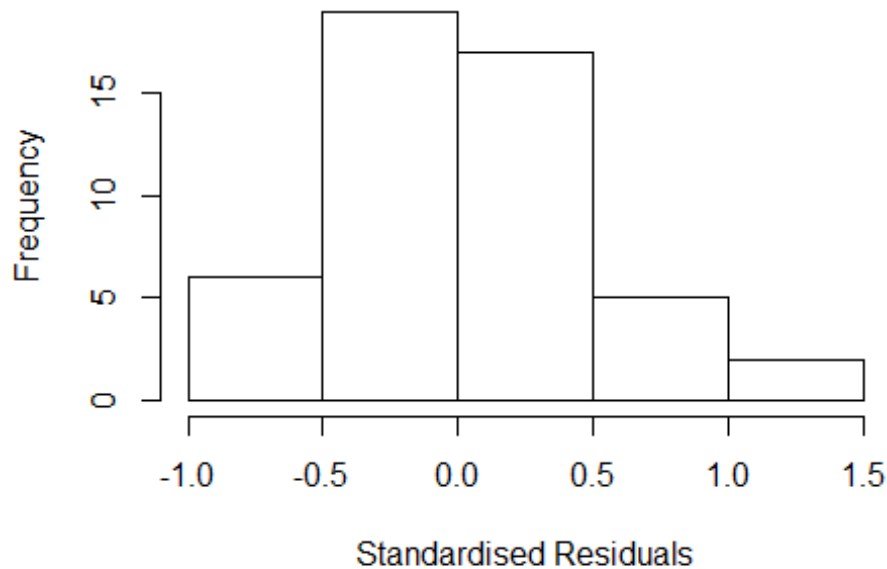
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.63e-01 -2.26e-01 -2.78e-16 2.81e-01 1.41e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5835 0.0279 56.65 < 2e-16 ***
## FirstAuthorFemale1 0.0086 0.2248 0.04 0.96973
## Year1997 -0.1785 0.0279 -6.39 4.1e-07 ***
## Year1998 -0.7040 0.2605 -2.70 0.01105 *
## Year1999 -0.9608 0.2380 -4.04 0.00033 ***
## Year2000 -1.0258 0.1245 -8.24 2.7e-09 ***
## Year2001 0.4440 0.2617 1.70 0.09986 .
## Year2002 -0.2504 0.2661 -0.94 0.35393
## Year2003 -0.7472 0.4036 -1.85 0.07364 .
## Year2004 -0.8840 0.1664 -5.31 8.7e-06 ***
## Year2005 -1.0158 0.2430 -4.18 0.00022 ***
## Year2006 0.1369 0.2265 0.60 0.54994
```

```

## Year2007          -0.3702      0.2595   -1.43   0.16379
## Year2008          -0.9161      0.1909   -4.80   3.8e-05 ***
## Year2009          -0.4348      0.3636   -1.20   0.24086
## Year2010          -0.5190      0.4841   -1.07   0.29194
## Year2011          -0.0425      0.0279   -1.52   0.13856
## Year2012          -0.6627      0.1441   -4.60   6.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.445, Adjusted R-squared:  0.141
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.535  0.921  0.978  0.941  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      2.04e-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 1.934e+16 1      1.391e+08
## Year              1.934e+16 16      3.228e+00

```

Residuals from last author



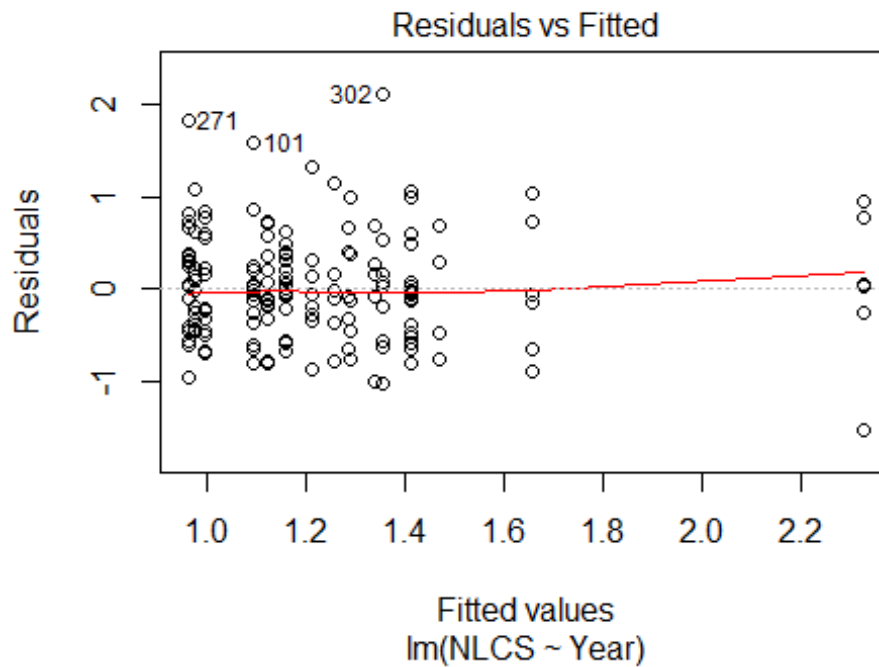
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.69e-01 -2.25e-01 6.59e-16 2.20e-01 1.38e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5835 0.0280 56.63 < 2e-16 ***
## LastAuthorFemale1 -0.3027 0.3497 -0.87 0.39344
## Year1997 -0.1785 0.0280 -6.38 4.1e-07 ***
## Year1998 -0.7040 0.2695 -2.61 0.01373 *
## Year1999 -0.8052 0.1986 -4.05 0.00031 ***
## Year2000 -1.0215 0.0564 -18.11 < 2e-16 ***
## Year2001 0.4521 0.2685 1.68 0.10230
## Year2002 -0.2324 0.2698 -0.86 0.39572
## Year2003 -0.6141 0.5548 -1.11 0.27685
## Year2004 -0.8840 0.1683 -5.25 1.0e-05 ***
## Year2005 -0.8602 0.2010 -4.28 0.00017 ***
## Year2006 0.4482 0.3509 1.28 0.21094
```

```

## Year2007          -0.3061      0.2128    -1.44   0.16031
## Year2008          -0.6081      0.3695    -1.65   0.10992
## Year2009          -0.4019      0.2810    -1.43   0.16267
## Year2010          -0.5190      0.5352    -0.97   0.33966
## Year2011          -0.0425      0.0280    -1.52   0.13870
## Year2012          -0.6620      0.1507    -4.39   0.00012 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.483, Adjusted R-squared:  0.199
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.401  0.902  0.971  0.919  0.993  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      2.04e-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: 49"
## [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
##   10   11   8   11   13   11   6   15   10   16   15   20   21   25   26
## 2011 2012
##   17   25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    6    5    6    5    8    5   14    6   12    5   13   13   15   20
## 2011 2012

```

```
## 11 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 6 5 6 5 8 5 12 6 11 2 11 11 14 16
## 2011 2012
## 8 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 = 15, df = 16, p-value = 0.5
```



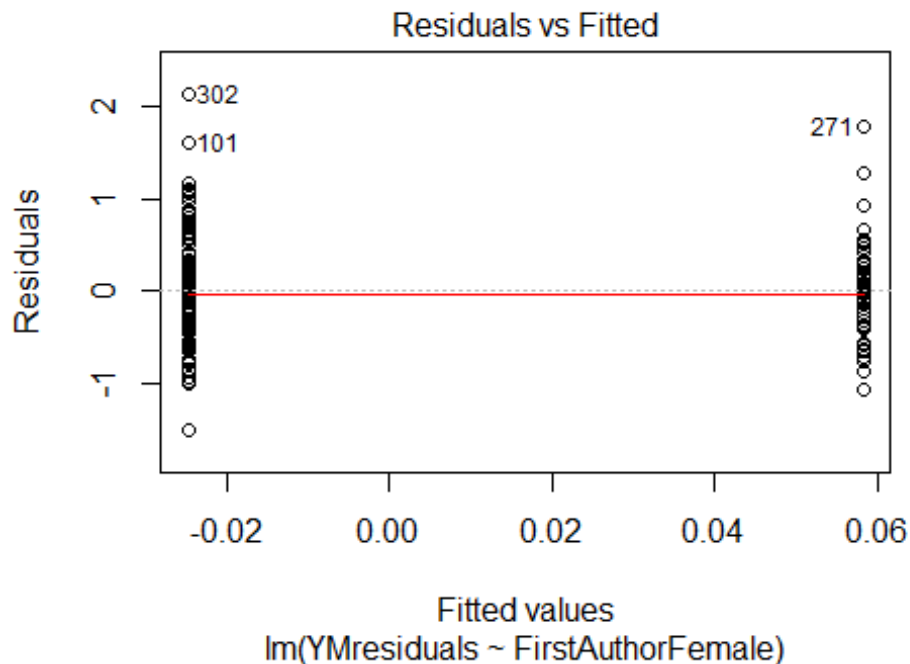
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.7, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 3.22370979547063"
## [1] "Male first author team size 2018 geometric mean: 2.21740488609733"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

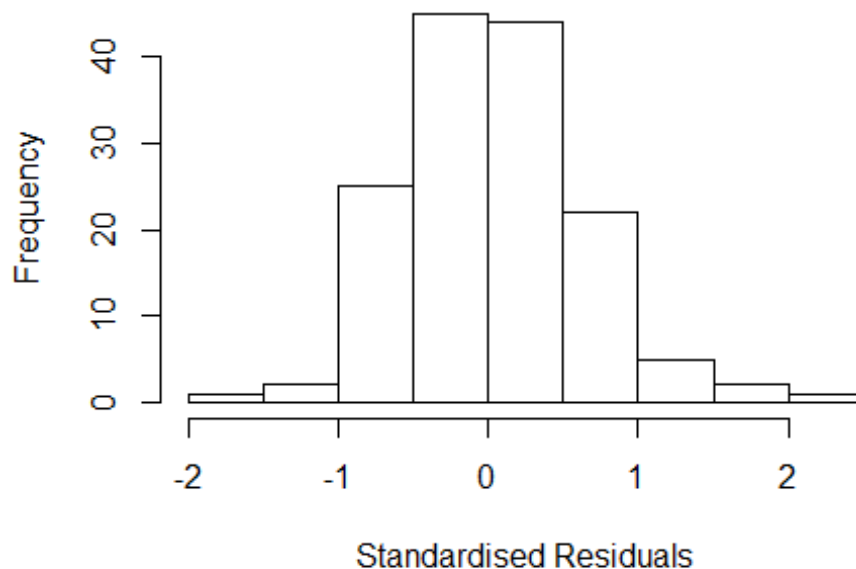
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 29, p-value = 0.09
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male last author team size 2018 geometric mean: 2.49505298836789"

## 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 = 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 3.323 1      1.823
## LastAuthorFemale  3.593 1      1.896
## UniqueAuthors    17.907 4      1.434
## Year              30.577 16     1.113
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.71508 -0.40327 0.00441 0.38201 2.07983
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1294 0.1761 6.41 2.7e-09 ***
## FirstAuthorFemale1 0.1606 0.1531 1.05 0.29606
## LastAuthorFemale1 -0.0642 0.1732 -0.37 0.71163
## UniqueAuthors2 -0.1896 0.1376 -1.38 0.17053
## UniqueAuthors3 -0.0810 0.1668 -0.49 0.62816
## UniqueAuthors4 0.1351 0.4309 0.31 0.75440
## UniqueAuthors5 -0.2616 0.3831 -0.68 0.49594
## Year1997 1.3786 0.3627 3.80 0.00022 ***
## Year1998 0.3174 0.2826 1.12 0.26343
## Year1999 0.5452 0.3840 1.42 0.15815
```

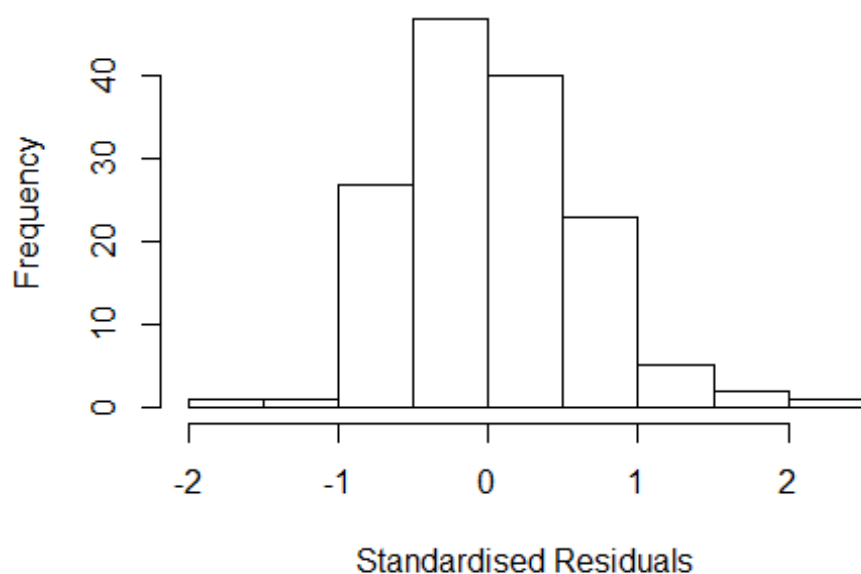


```

## Year2000          0.2231      0.3035      0.74  0.46359
## Year2001          0.0191      0.2475      0.08  0.93871
## Year2002          0.4391      0.3592      1.22  0.22384
## Year2003         -0.1791      0.2523     -0.71  0.47904
## Year2004          0.1173      0.2977      0.39  0.69427
## Year2005         -0.0927      0.2408     -0.39  0.70077
## Year2006         -0.4919      0.2907     -1.69  0.09312 .
## Year2007         -0.1313      0.2433     -0.54  0.59034
## Year2008          0.0225      0.1899      0.12  0.90578
## Year2009         -0.0405      0.2091     -0.19  0.84659
## Year2010         -0.1039      0.2705     -0.38  0.70155
## Year2011          0.1858      0.4029      0.46  0.64540
## Year2012          0.1915      0.2221      0.86  0.39025
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.566
## Multiple R-squared:  0.263, Adjusted R-squared:  0.133
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 124 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.148  0.854  0.939  0.892  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      6.80e-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 2.775 1      1.666
## LastAuthorFemale  1.871 1      1.368
## Year              3.279 16      1.038

```

Residuals from first and last author



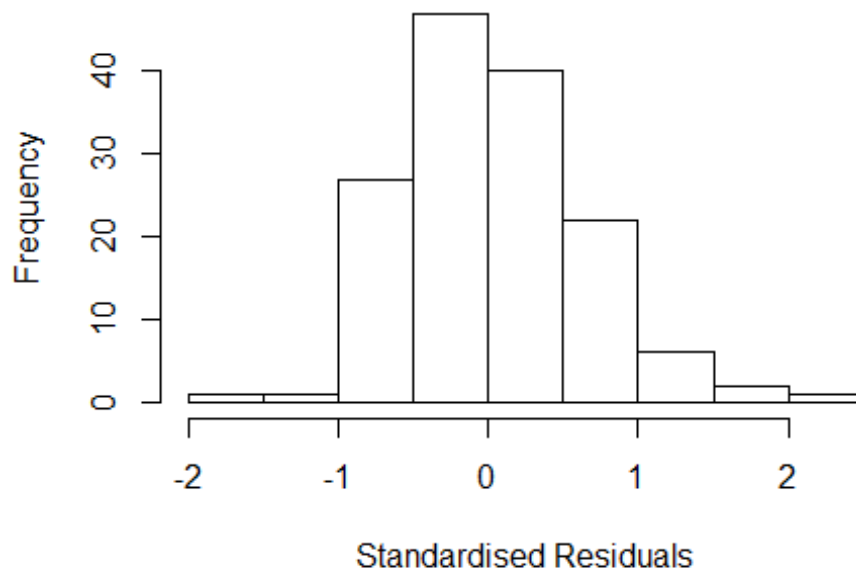
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5971 -0.4082 -0.0163  0.3486  2.2448
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.077278   0.158747   6.79 3.8e-10 ***
## FirstAuthorFemale1 0.104074   0.143523   0.73 0.46969
## LastAuthorFemale1 0.024252   0.141608   0.17 0.86429
## Year1997        1.312847   0.379329   3.46 0.00073 ***
## Year1998        0.302266   0.304961   0.99 0.32348
## Year1999        0.537027   0.361445   1.49 0.13980
## Year2000        0.190696   0.298249   0.64 0.52372
## Year2001        0.039498   0.242346   0.16 0.87079
## Year2002        0.401801   0.316840   1.27 0.20705
## Year2003       -0.127054   0.215601  -0.59 0.55670
## Year2004        0.134335   0.287168   0.47 0.64073
## Year2005       -0.143096   0.212843  -0.67 0.50260
```

```

## Year2006          -0.416404    0.203957    -2.04    0.04324 *
## Year2007          -0.197896    0.226440    -0.87    0.38379
## Year2008          -0.000218    0.182910     0.00    0.99905
## Year2009          -0.059176    0.182807    -0.32    0.74669
## Year2010          -0.159818    0.246977    -0.65    0.51873
## Year2011           0.119694    0.371780     0.32    0.74802
## Year2012           0.217612    0.206628     1.05    0.29425
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.232, Adjusted R-squared:  0.124
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.138  0.880   0.944   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      6.80e-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 2.167 1      1.472
## Year              2.167 16      1.024

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.602 -0.400 -0.014 0.353 2.264
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.07928 0.15692 6.88 2.4e-10 ***
## FirstAuthorFemale1 0.11391 0.12645 0.90 0.36935
## Year1997 1.31536 0.37958 3.47 0.00072 ***
## Year1998 0.30125 0.30400 0.99 0.32356
## Year1999 0.53266 0.35976 1.48 0.14116
## Year2000 0.18727 0.29970 0.62 0.53316
## Year2001 0.04060 0.24294 0.17 0.86753
## Year2002 0.40012 0.31666 1.26 0.20868
## Year2003 -0.13455 0.20777 -0.65 0.51841
## Year2004 0.13120 0.28619 0.46 0.64741
## Year2005 -0.14090 0.21110 -0.67 0.50567
## Year2006 -0.40628 0.19135 -2.12 0.03564 *
```

```

## Year2007          -0.19471    0.22569   -0.86  0.38989
## Year2008           0.00133    0.18242    0.01  0.99421
## Year2009          -0.05930    0.18169   -0.33  0.74465
## Year2010          -0.15891    0.24595   -0.65  0.51935
## Year2011           0.12249    0.37692    0.32  0.74573
## Year2012           0.21379    0.20169    1.06  0.29112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.598
## Multiple R-squared:  0.233, Adjusted R-squared:  0.132
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.121  0.878  0.941  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      6.80e-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 1.386 1      1.177
## Year              1.386 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
## -1.6397 -0.3982 -0.0322  0.3530  2.1942

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.13008    0.15548   7.27 3.1e-11 ***
## LastAuthorFemale1 0.07212    0.12437   0.58 0.56302
## Year1997        1.30259    0.36577   3.56 0.00052 ***
## Year1998        0.25001    0.30325   0.82 0.41120
## Year1999        0.50012    0.37892   1.32 0.18922
## Year2000        0.15440    0.28295   0.55 0.58622
## Year2001        0.00246    0.24253   0.01 0.99192
## Year2002        0.34917    0.31567   1.11 0.27072
## Year2003       -0.14603    0.22829  -0.64 0.52352
## Year2004        0.08090    0.28551   0.28 0.77735
## Year2005       -0.20592    0.20423  -1.01 0.31520
## Year2006       -0.49314    0.19659  -2.51 0.01337 *
## Year2007       -0.22923    0.22526  -1.02 0.31075
## Year2008       -0.01092    0.19377  -0.06 0.95513
## Year2009       -0.05797    0.19045  -0.30 0.76133
## Year2010       -0.21127    0.24648  -0.86 0.39295
## Year2011        0.06964    0.36274   0.19 0.84805
## Year2012        0.20486    0.21770   0.94 0.34847
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.601
## Multiple R-squared:  0.232, Adjusted R-squared:  0.13
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 131 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.154 0.878  0.950  0.906  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      6.80e-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: 147"
## [1] ""

```

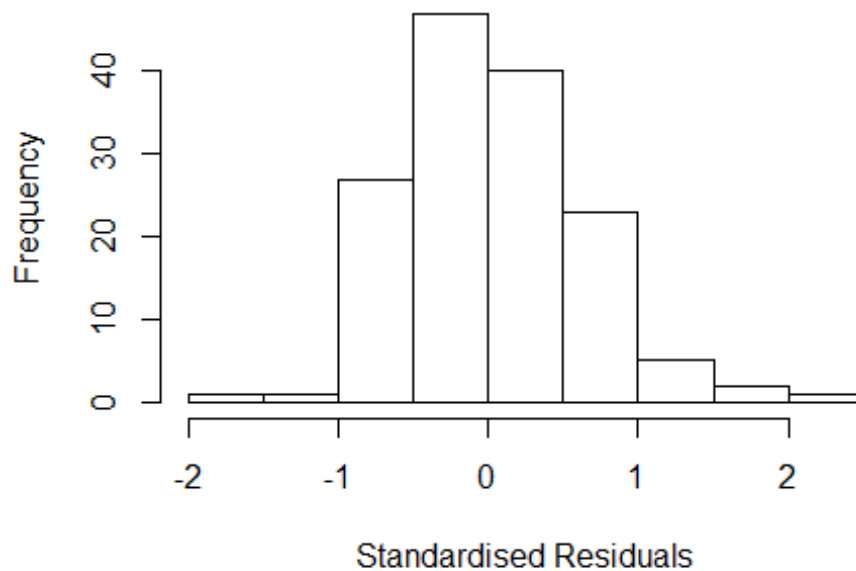
```

## [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
##    2    4    2    6    3    8    2    1    2    1    8    6    6    4    12
## 2011 2012
##   11    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    0    3    2    2    0    1    1    1    6    2    4    2    6
## 2011 2012
##    3    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    0    3    2    2    0    0    1    0    4    1    3    1    5
## 2011 2012
##    2    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 3.13016916014657"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 3.10369114783072"
## [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"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 13.247  1          3.640
## LastAuthorFemale  2.357  1          1.535
## Year              25.681 12          1.145
##
## [1] "List of 0 outliers with 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"
```



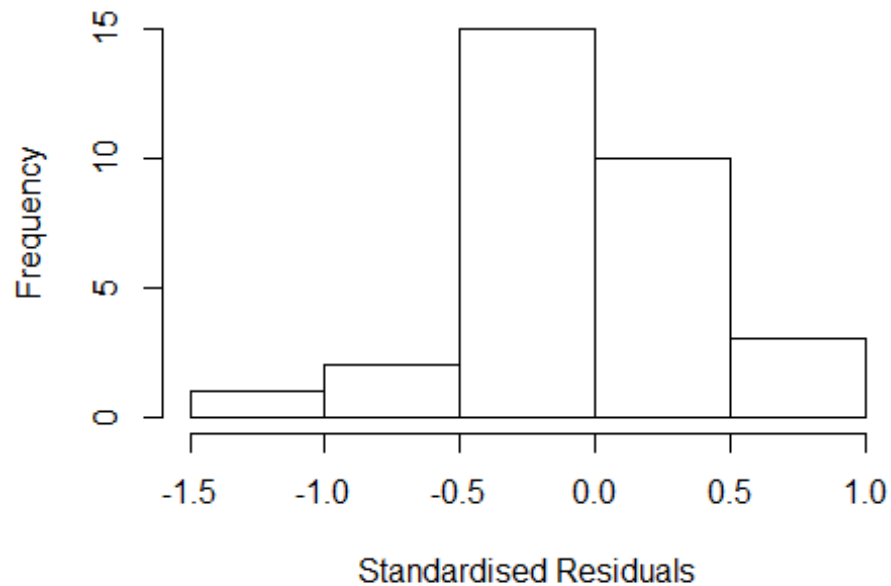
```

## Residuals:
##      Min      1Q    Median      3Q      Max
## -1.14e+00 -1.85e-01 -3.05e-16  3.34e-01  6.96e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.97373    0.13731    7.09 2.6e-06 ***
## FirstAuthorFemale1 0.75655    0.26787    2.82  0.012 *
## LastAuthorFemale1 -0.50018    0.32680   -1.53  0.145
## Year1997         0.27827    0.36509    0.76  0.457
## Year1999        -0.07713    0.23726   -0.33  0.749
## Year2000         0.28427    0.66182    0.43  0.673
## Year2001        -0.00241    0.32966   -0.01  0.994
## Year2004         0.20927    0.13731    1.52  0.147
## Year2006        -0.21923    0.26408   -0.83  0.419
## Year2007        -0.07273    0.13731   -0.53  0.604
## Year2008         0.19779    0.22809    0.87  0.399
## Year2009         0.25527    0.13731    1.86  0.081 .
## Year2010         0.67460    0.30250    2.23  0.040 *
## Year2011        -0.20123    0.14282   -1.41  0.178
## Year2012        -0.06443    0.24583   -0.26  0.797
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.402, Adjusted R-squared:  -0.122
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 21 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.719  0.936  0.967  0.945  0.978  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      3.23e-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"

```

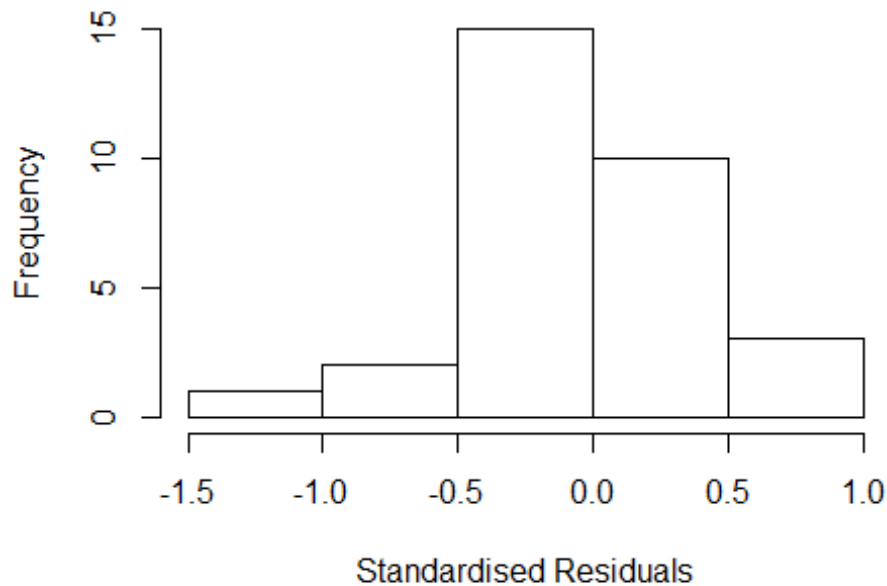
```
## Warning in lmrob.S(x, y, control = control, mf = mf): S refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps
## Warning in lmrob.S(x, y, control = control, mf = mf): S refinements did
not
## converge (to refine.tol=1e-07) in 1000 (= k.max) steps
```

Residuals from first and last author



```
## [1] "Regression 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 4.133 1      2.033
## Year            4.133 12      1.061
```

Residuals from last author



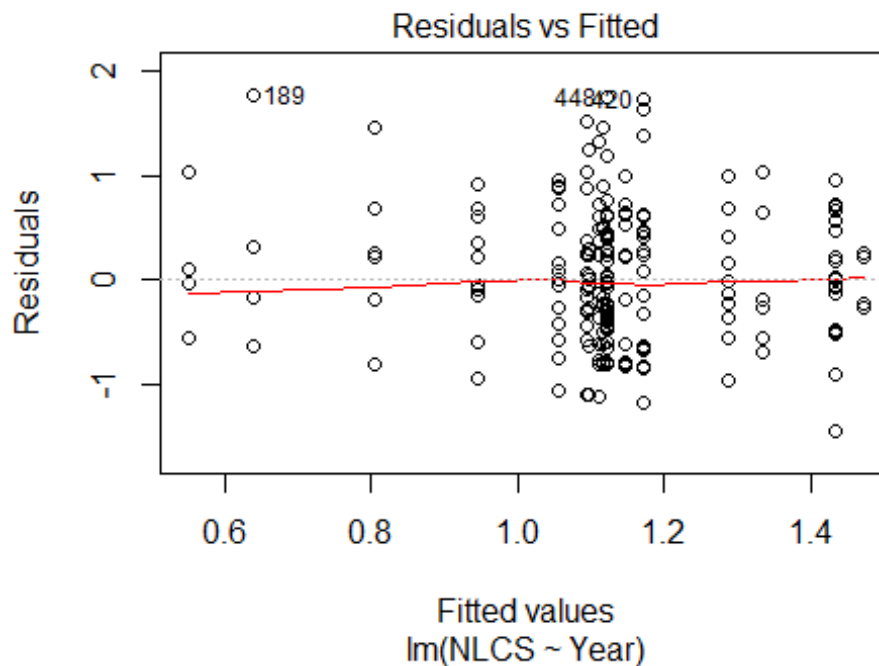
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.14e+00 -2.80e-01 -4.16e-16 3.31e-01 9.44e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3520 0.3235 4.18 0.00063 ***
## LastAuthorFemale1 -0.3781 0.2653 -1.43 0.17223
## Year1997 -0.1000 0.4666 -0.21 0.83284
## Year1999 -0.4553 0.3767 -1.21 0.24332
## Year2000 -0.0940 0.7131 -0.13 0.89668
## Year2001 -0.0634 0.3521 -0.18 0.85911
## Year2004 -0.1690 0.3235 -0.52 0.60810
## Year2006 -0.4563 0.5707 -0.80 0.43498
## Year2007 -0.4510 0.3235 -1.39 0.18120
## Year2008 -0.2212 0.3620 -0.61 0.54929
## Year2009 -0.1230 0.3235 -0.38 0.70847
## Year2010 0.2928 0.4203 0.70 0.49546
```

```

## Year2011          -0.5795      0.3259   -1.78  0.09323 .
## Year2012          -0.4424      0.3821   -1.16  0.26297
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.643
## Multiple R-squared:  0.298, Adjusted R-squared:  -0.239
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 20 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.735  0.928  0.961  0.935  0.975  0.995
## Algorithmic parameters:
##           tuning.chi             bb           tuning.psi           refine.tol
##           1.55e+00             5.00e-01           4.69e+00           1.00e-07
##           rel.tol             solve.tol           eps.outlier           eps.x
##           1.00e-07             1.00e-07           3.23e-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: 31"
## [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
##   15   17   18    9   17   28   20   15   15   13   25   27   33   26   36
## 2011 2012
##   43   40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11    6   13    5    4   16   11   11    6    8   10   12   15   11   25
## 2011 2012
##   27   19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10    6    8    2    4   16   11    8    5    6    7   10   11    9   21

```

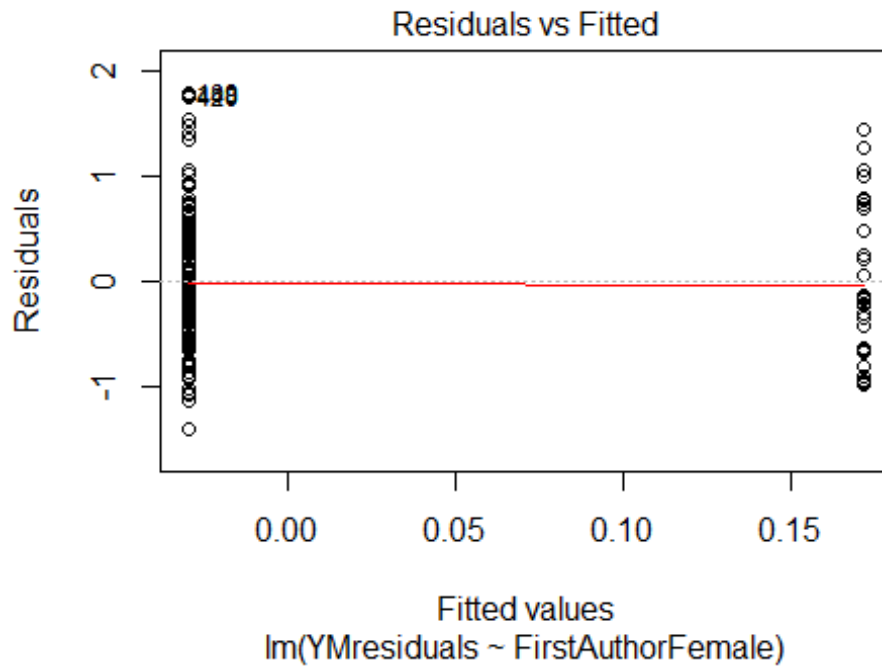
```
## 2011 2012
## 21 12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 10, df = 16, p-value = 0.9
```



```
##
## 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: 4"
## [1] "Male first author team size 2018 geometric mean: 2.1689435423954"

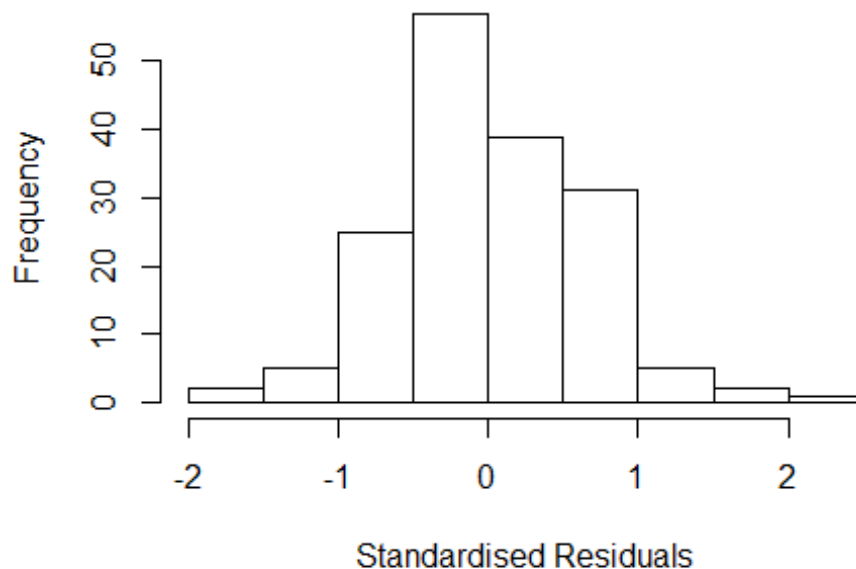
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8.5, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 2.29304697369141"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF^(1/(2*Df))
## FirstAuthorFemale	1.971	1	1.404
## LastAuthorFemale	2.271	1	1.507
## UniqueAuthors	11.540	4	1.358
## Year	19.390	16	1.097

Residuals from first and last author and team size



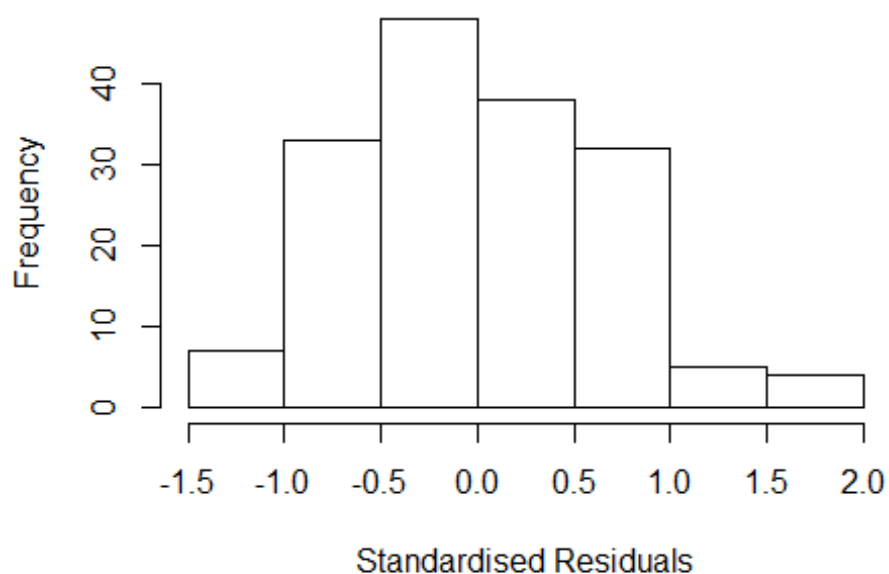
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5703 -0.3836 -0.0232 0.4770 2.1863
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.87265 0.23764 3.67 0.00034 ***
## FirstAuthorFemale1 0.27184 0.14624 1.86 0.06510 .
## LastAuthorFemale1 -0.26492 0.14129 -1.87 0.06282 .
## UniqueAuthors2 0.32590 0.13130 2.48 0.01421 *
## UniqueAuthors3 0.33833 0.15365 2.20 0.02926 *
## UniqueAuthors4 0.65941 0.23355 2.82 0.00543 **
## UniqueAuthors5 0.74574 0.24570 3.04 0.00285 **
## Year1997 0.27408 0.32379 0.85 0.39869
## Year1998 0.00993 0.34606 0.03 0.97716
## Year1999 0.01540 0.36934 0.04 0.96680
```

```

## Year2000          0.40387    0.30085    1.34  0.18157
## Year2001          0.00737    0.27519    0.03  0.97866
## Year2002          0.01451    0.29235    0.05  0.96047
## Year2003         -0.07545    0.30032   -0.25  0.80200
## Year2004         -0.65896    0.38290   -1.72  0.08740 .
## Year2005         -0.88582    0.46946   -1.89  0.06119 .
## Year2006         -0.12585    0.32636   -0.39  0.70034
## Year2007         -0.08758    0.35989   -0.24  0.80808
## Year2008         -0.17035    0.28733   -0.59  0.55419
## Year2009         -0.35748    0.32371   -1.10  0.27129
## Year2010         -0.04804    0.33537   -0.14  0.88629
## Year2011         -0.31084    0.25826   -1.20  0.23072
## Year2012          0.05353    0.29404    0.18  0.85581
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.608
## Multiple R-squared:  0.216, Adjusted R-squared:  0.0968
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 153 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.169  0.892  0.944  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      5.99e-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 1.529 1      1.237
## LastAuthorFemale  1.333 1      1.155
## Year              2.034 16      1.022

```


Residuals from first and last author



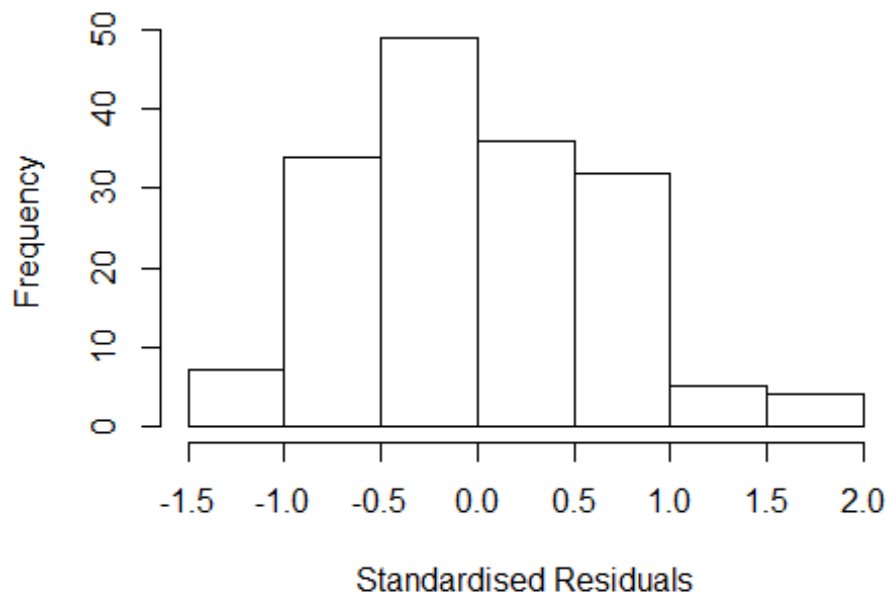
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1957 -0.4740 -0.0544 0.4709 1.8108
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.89674 0.21817 4.11 6.5e-05 ***
## FirstAuthorFemale1 0.33269 0.16828 1.98 0.05 *
## LastAuthorFemale1 -0.18570 0.15601 -1.19 0.24
## Year1997 0.40610 0.35781 1.13 0.26
## Year1998 0.15673 0.32700 0.48 0.63
## Year1999 0.15426 0.47758 0.32 0.75
## Year2000 0.66857 0.25673 2.60 0.01 *
## Year2001 0.11513 0.27343 0.42 0.67
## Year2002 0.23101 0.27418 0.84 0.40
## Year2003 0.21468 0.28349 0.76 0.45
## Year2004 -0.30253 0.45284 -0.67 0.51
## Year2005 -0.44359 0.42365 -1.05 0.30
```

```

## Year2006          0.20542      0.28924      0.71      0.48
## Year2007          0.05847      0.32579      0.18      0.86
## Year2008          0.04262      0.30668      0.14      0.89
## Year2009         -0.00657      0.32618     -0.02      0.98
## Year2010          0.26907      0.29037      0.93      0.36
## Year2011          0.05161      0.23241      0.22      0.82
## Year2012          0.29898      0.30138      0.99      0.32
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.699
## Multiple R-squared:  0.0988, Adjusted R-squared:  -0.0108
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 155 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.482  0.894  0.947  0.922  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      5.99e-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 1.518 1      1.232
## Year              1.518 16      1.013

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.191 -0.489 -0.033 0.489 1.805
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8765 0.2147 4.08 7.2e-05 ***
## FirstAuthorFemale1 0.3095 0.1771 1.75 0.083 .
## Year1997 0.4267 0.3557 1.20 0.232
## Year1998 0.1769 0.3245 0.55 0.586
## Year1999 0.1745 0.4756 0.37 0.714
## Year2000 0.5960 0.2487 2.40 0.018 *
## Year2001 0.1396 0.2703 0.52 0.606
## Year2002 0.2512 0.2713 0.93 0.356
## Year2003 0.2367 0.2819 0.84 0.402
## Year2004 -0.2800 0.4531 -0.62 0.537
## Year2005 -0.4175 0.4206 -0.99 0.322
## Year2006 0.2257 0.2866 0.79 0.432
```

```

## Year2007          0.0483      0.3262      0.15      0.883
## Year2008          0.0341      0.3161      0.11      0.914
## Year2009          0.0237      0.3224      0.07      0.941
## Year2010          0.2665      0.2890      0.92      0.358
## Year2011          0.0325      0.2311      0.14      0.888
## Year2012          0.3150      0.3033      1.04      0.301
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.703
## Multiple R-squared:  0.0935, Adjusted R-squared:  -0.00995
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 152 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.489  0.893  0.944  0.920  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      5.99e-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 1.321 1      1.149
## 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.374 -0.487 -0.027  0.444  1.803

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9290    0.2210   4.20 4.5e-05 ***
## LastAuthorFemale1 -0.1205    0.1360  -0.89   0.38
## Year1997        0.3742    0.3605   1.04   0.30
## Year1998        0.1244    0.3284   0.38   0.71
## Year1999        0.1220    0.4782   0.26   0.80
## Year2000        0.6037    0.2567   2.35   0.02 *
## Year2001        0.1359    0.2837   0.48   0.63
## Year2002        0.1986    0.2759   0.72   0.47
## Year2003        0.2042    0.3018   0.68   0.50
## Year2004       -0.3324    0.4591  -0.72   0.47
## Year2005       -0.4103    0.4012  -1.02   0.31
## Year2006        0.1732    0.2916   0.59   0.55
## Year2007        0.0476    0.3175   0.15   0.88
## Year2008        0.0772    0.3297   0.23   0.82
## Year2009        0.0980    0.3232   0.30   0.76
## Year2010        0.2515    0.3009   0.84   0.40
## Year2011        0.0670    0.2349   0.29   0.78
## Year2012        0.4454    0.3024   1.47   0.14
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.703
## Multiple R-squared:  0.0755, Adjusted R-squared:  -0.03
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 156 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.490  0.901   0.951   0.921   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      5.99e-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: 167"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012
##    6    3    2    2   16    3    1    3    1    8    7    9    7    3    1
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012
##    3    2    2    1   10    3    0    3    1    8    7    9    3    3    1
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012
##    1    2    2    1    9    3    0    2    1    8    7    7    3    3    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 3"

## 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 = 1, p-value = NA
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 3"

## 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 = 1, 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"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate

## [1] "Regression 2: First author gender, Last 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 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"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate

## [1] "Sample size for the above analysis: 50"
## [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]"
## < table of extent 0 >
## < table of extent 0 >
## < table of extent 0 >
## [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 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: 0"
## [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
##    3    9    6    4   10    5    8    7    9   10   14   15   11   13    9
## 2011 2012
##   11    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    2    3    8    4    7    7    8    9   10   12    6   13    7
## 2011 2012
##    8    1
##

```

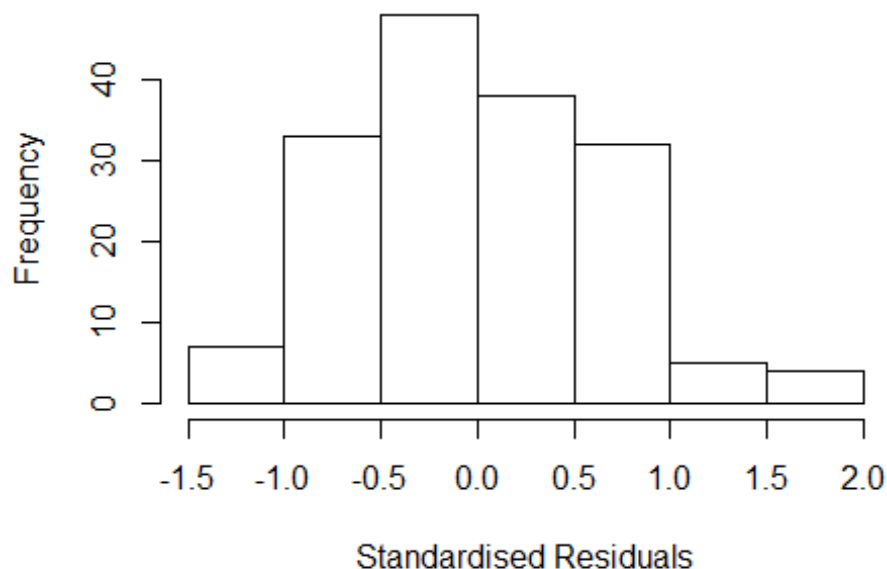
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    1    3    8    4    7    4    7    8   10    8    5   13    6
## 2011 2012
##    4    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.82842712474619"
## [1] "Male first author team size 2018 geometric mean: 2.1689435423954"

## 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 = 6.5, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.28942848510666"
## [1] "Male last author team size 2018 geometric mean: 2.37841423000544"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

Residuals from last author

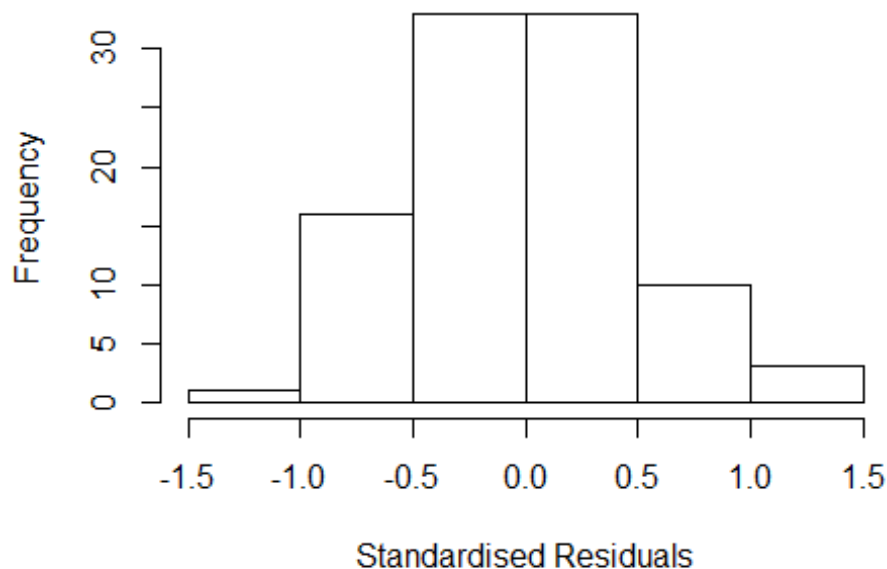


```
##
## Wilcoxon rank sum test with continuity correction
```



```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5, 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 5.049e+00  1      2.247
## LastAuthorFemale  1.709e+00  1      1.307
## UniqueAuthors    5.217e+14  2    4779.272
## Year              1.188e+15 16      2.959
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.06655 -0.36742 -0.00416 0.36656 1.25424
```

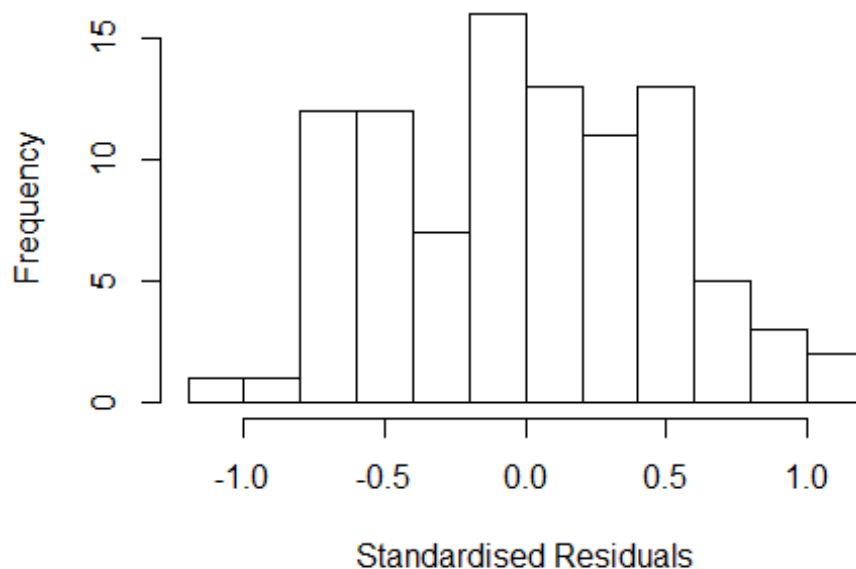
```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9936    0.4497    2.21  0.030 *
## FirstAuthorFemale1 -0.1200    0.1554   -0.77  0.442
## LastAuthorFemale1  0.1455    0.1692    0.86  0.393
## UniqueAuthors2    0.1550    0.1385    1.12  0.267
## UniqueAuthors3    0.3626    0.1688    2.15  0.035 *
## Year1997         -0.0433    0.5129   -0.08  0.933
## Year1998          0.2844    0.4497    0.63  0.529
## Year1999         -0.1436    0.5398   -0.27  0.791
## Year2000         -0.5094    0.4557   -1.12  0.267
## Year2001          0.3589    0.4730    0.76  0.450
## Year2002          0.3161    0.5103    0.62  0.537
## Year2003          0.0729    0.5776    0.13  0.900
## Year2004         -0.6265    0.4752   -1.32  0.191
## Year2005          0.0346    0.4715    0.07  0.942
## Year2006         -0.3793    0.4568   -0.83  0.409
## Year2007         -0.0724    0.5101   -0.14  0.888
## Year2008         -0.1163    0.4545   -0.26  0.799
## Year2009         -0.1823    0.4592   -0.40  0.692
## Year2010         -0.1078    0.4999   -0.22  0.830
## Year2011         -0.8471    0.4758   -1.78  0.079 .
## Year2012         -0.5006    0.4527   -1.11  0.272
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.574
## Multiple R-squared:  0.289, Adjusted R-squared:  0.0988
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 88 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.613  0.906  0.962  0.935  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      1.04e-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 2.076 1          1.441
## LastAuthorFemale  1.748 1          1.322
## Year              5.666 16          1.056
```

Residuals from first and last author



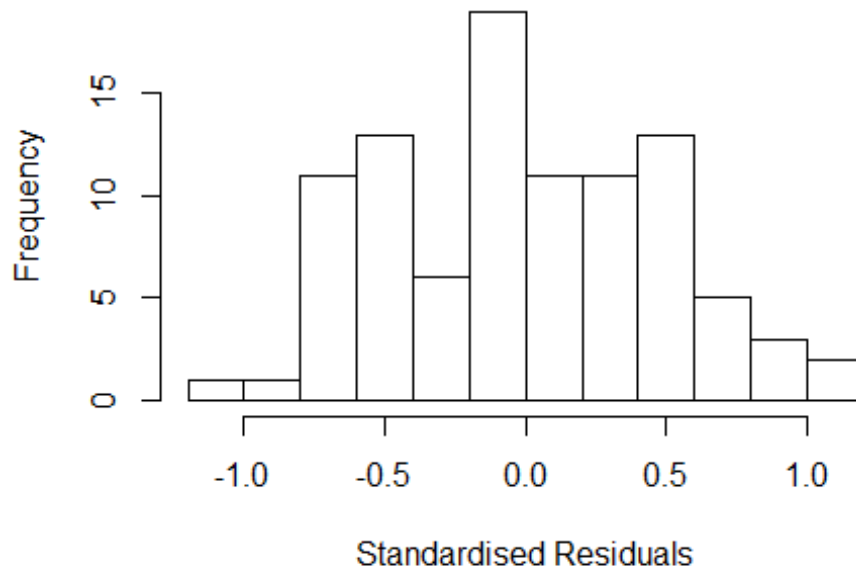
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.08e+00 -4.06e-01 -9.30e-16  3.79e-01  1.07e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0466      0.3844   2.72   0.008 **
## FirstAuthorFemale1 -0.0197      0.1489  -0.13   0.895
## LastAuthorFemale1  0.0848      0.1888   0.45   0.655
## Year1997          -0.0851      0.4727  -0.18   0.858
```

```

## Year1998          0.2314      0.3844      0.60      0.549
## Year1999          -0.0942     0.4771     -0.20     0.844
## Year2000          -0.4929     0.4059     -1.21     0.228
## Year2001          0.2939     0.4141      0.71     0.480
## Year2002          0.3681     0.4621      0.80     0.428
## Year2003          0.0374     0.5376      0.07     0.945
## Year2004          -0.6342     0.4327     -1.47     0.147
## Year2005          0.1307     0.4050      0.32     0.748
## Year2006          -0.3720     0.4010     -0.93     0.357
## Year2007          -0.0501     0.4450     -0.11     0.911
## Year2008          -0.0315     0.4141     -0.08     0.940
## Year2009          -0.0943     0.3866     -0.24     0.808
## Year2010          -0.1055     0.4942     -0.21     0.831
## Year2011          -0.7837     0.4172     -1.88     0.064 .
## Year2012          -0.3986     0.3844     -1.04     0.303
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.263, Adjusted R-squared:  0.0907
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.644  0.876  0.944  0.913  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          1.04e-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 1.455 1          1.206
## Year              1.455 16          1.012

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.0597 -0.4093 -0.0183 0.3622 1.1362
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03680 0.37188 2.79 0.0067 **
## FirstAuthorFemale1 0.02045 0.12844 0.16 0.8739
## Year1997 -0.08375 0.45217 -0.19 0.8535
## Year1998 0.24120 0.37188 0.65 0.5185
## Year1999 -0.09157 0.46498 -0.20 0.8444
## Year2000 -0.47648 0.39448 -1.21 0.2307
## Year2001 0.31751 0.39641 0.80 0.4256
## Year2002 0.38192 0.44194 0.86 0.3901
## Year2003 0.02286 0.51959 0.04 0.9650
## Year2004 -0.60771 0.41257 -1.47 0.1448
## Year2005 0.15683 0.39118 0.40 0.6896
## Year2006 -0.36565 0.39141 -0.93 0.3531
```

```

## Year2007      -0.00182    0.41758    0.00    0.9965
## Year2008      -0.01782    0.39795   -0.04    0.9644
## Year2009      -0.10243    0.37743   -0.27    0.7868
## Year2010      -0.08059    0.46250   -0.17    0.8621
## Year2011      -0.75828    0.39927   -1.90    0.0612 .
## Year2012      -0.38880    0.37188   -1.05    0.2990
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.255, Adjusted R-squared:  0.0928
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.703  0.897  0.952  0.933  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      1.04e-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 3.832 1      1.957
## Year      3.832 16      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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.08446 -0.40717 -0.00428  0.36367  1.07685

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.0353    0.3510    2.95  0.0042 **
## LastAuthorFemale1 0.0734    0.1703    0.43  0.6677
## Year1997       -0.0775    0.4459   -0.17  0.8625
## Year1998        0.2427    0.3510    0.69  0.4913
## Year1999       -0.0827    0.4523   -0.18  0.8554
## Year2000       -0.4838    0.3832   -1.26  0.2105
## Year2001        0.3028    0.3915    0.77  0.4415
## Year2002        0.3741    0.4524    0.83  0.4109
## Year2003        0.0491    0.5174    0.09  0.9246
## Year2004       -0.6251    0.4062   -1.54  0.1279
## Year2005        0.1422    0.3749    0.38  0.7055
## Year2006       -0.3643    0.3830   -0.95  0.3444
## Year2007       -0.0385    0.4135   -0.09  0.9260
## Year2008       -0.0249    0.3924   -0.06  0.9496
## Year2009       -0.0945    0.3828   -0.25  0.8056
## Year2010       -0.1045    0.4862   -0.21  0.8304
## Year2011       -0.7748    0.3944   -1.96  0.0530 .
## Year2012       -0.3873    0.3510   -1.10  0.2732
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.263, Adjusted R-squared:  0.102
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.642  0.872  0.942  0.913  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      1.04e-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: 96"
## [1] ""

```

```

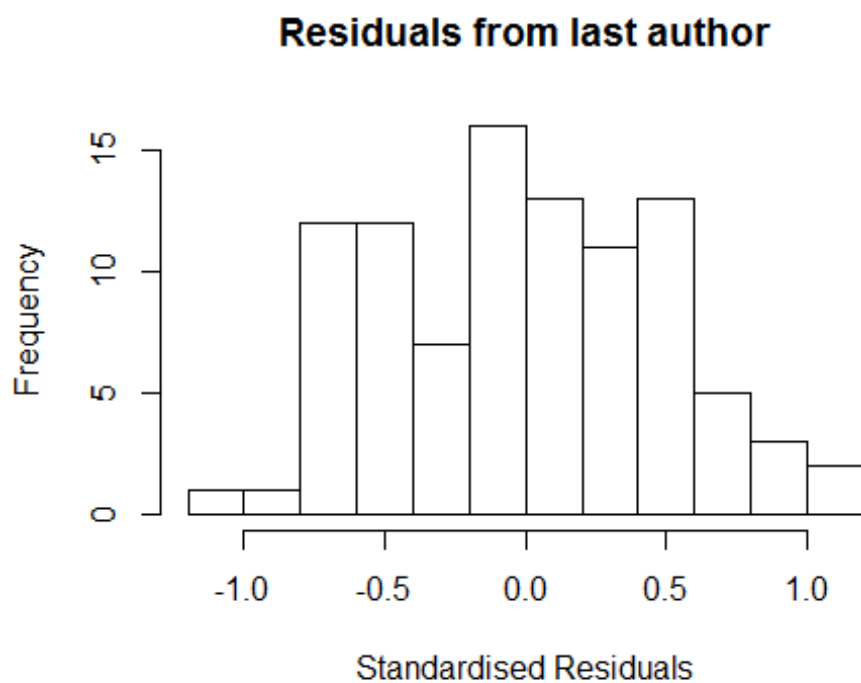
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1803"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 24 16 7 5 17 13 17 11 13 17 12 17 14 13
## 2011 2012
## 15 9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 13 5 6 1 9 4 10 8 8 14 10 9 13 10
## 2011 2012
## 9 3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 13 4 6 1 8 4 9 6 7 13 8 8 11 10
## 2011 2012
## 7 2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 2.21336383940064"

## 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 = 5, p-value = 0.8
## 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: 2.37956557896877"

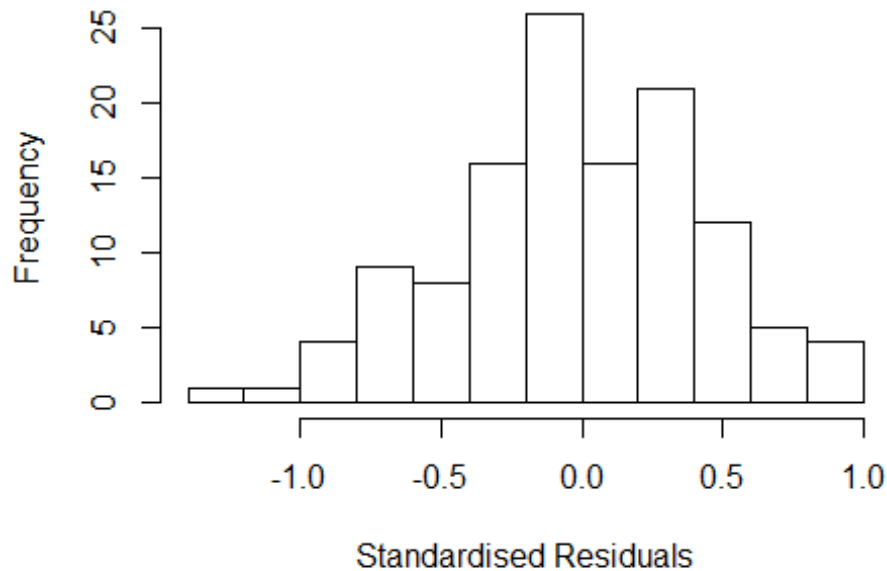
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4, p-value = 0.5
## 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  10.145  1          3.185
## LastAuthorFemale   9.367  1          3.061
## UniqueAuthors     291.723  3          2.575
## Year               3193.124 16          1.287
```

Residuals from first and last author and team size



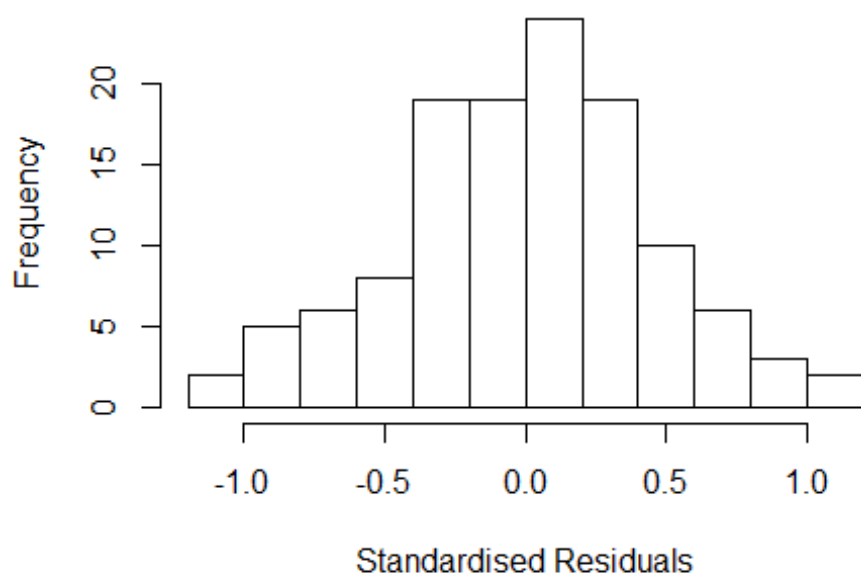
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22269 -0.26857 -0.00694 0.28677 0.96071
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3552 0.1310 10.34 < 2e-16 ***
## FirstAuthorFemale1 0.0175 0.1584 0.11 0.9121
## LastAuthorFemale1 0.1609 0.1624 0.99 0.3242
## UniqueAuthors2 0.1051 0.1020 1.03 0.3056
## UniqueAuthors3 0.1028 0.2048 0.50 0.6166
## UniqueAuthors4 -0.1427 0.2557 -0.56 0.5781
## Year1997 -0.5336 0.1945 -2.74 0.0072 **
## Year1998 -0.0959 0.1837 -0.52 0.6030
## Year1999 -0.5157 0.3055 -1.69 0.0945 .
## Year2000 -0.7223 0.1583 -4.56 1.4e-05 ***
```

```

## Year2001          -0.4146      0.1865    -2.22    0.0284 *
## Year2002          -0.2030      0.3396    -0.60    0.5513
## Year2003          -0.2003      0.1911    -1.05    0.2972
## Year2004          -0.2376      0.2563    -0.93    0.3561
## Year2005          -0.0122      0.2158    -0.06    0.9549
## Year2006          -0.4725      0.2003    -2.36    0.0202 *
## Year2007          -0.7793      0.1893    -4.12    7.9e-05 ***
## Year2008           0.0364      0.2185     0.17    0.8682
## Year2009          -0.1766      0.1799    -0.98    0.3286
## Year2010          -0.4289      0.2881    -1.49    0.1397
## Year2011          -0.2971      0.2674    -1.11    0.2691
## Year2012           0.2818      0.2338     1.21    0.2310
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.246, Adjusted R-squared:  0.0896
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.512  0.887   0.959   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      8.13e-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  7.045  1      2.654
## LastAuthorFemale   2.753  1      1.659
## Year              11.286 16      1.079

```

Residuals from first and last author



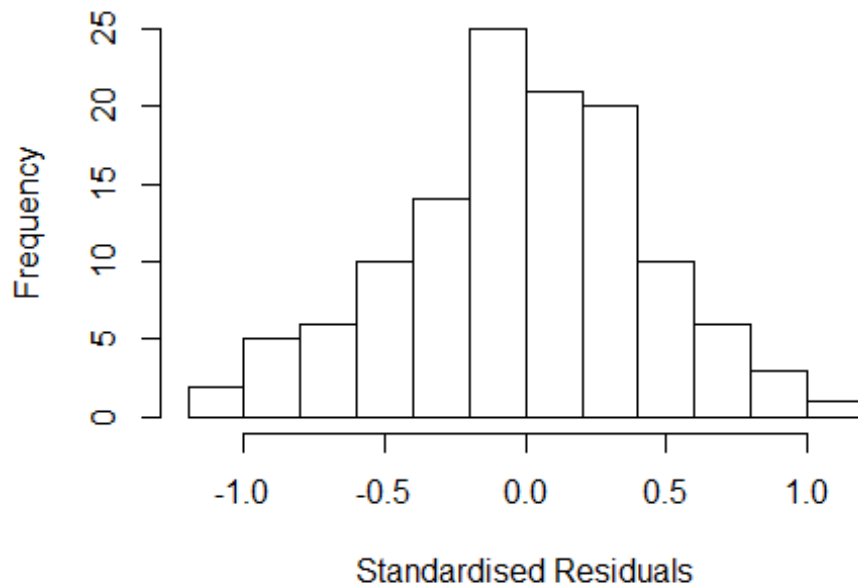
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1587 -0.2920  0.0115  0.2929  1.0167
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3695     0.1394   9.83 < 2e-16 ***
## FirstAuthorFemale1  0.0580     0.1432   0.40  0.68635
## LastAuthorFemale1  0.1406     0.1530   0.92  0.36012
## Year1997          -0.5124     0.2119  -2.42  0.01735 *
## Year1998          -0.0988     0.1835  -0.54  0.59144
## Year1999          -0.4308     0.3075  -1.40  0.16426
## Year2000          -0.6315     0.1394  -4.53  1.6e-05 ***
## Year2001          -0.3777     0.1807  -2.09  0.03906 *
## Year2002          -0.1559     0.3507  -0.44  0.65753
## Year2003          -0.1681     0.1972  -0.85  0.39617
## Year2004          -0.2108     0.2441  -0.86  0.38976
## Year2005           0.0398     0.1964   0.20  0.83993
```

```

## Year2006          -0.4250      0.1953   -2.18  0.03184 *
## Year2007          -0.7505      0.1964   -3.82  0.00023 ***
## Year2008           0.0824      0.2364    0.35  0.72800
## Year2009          -0.1367      0.1910   -0.72  0.47560
## Year2010          -0.4012      0.2760   -1.45  0.14903
## Year2011          -0.2328      0.2630   -0.89  0.37797
## Year2012           0.2285      0.1617    1.41  0.16055
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.477
## Multiple R-squared:  0.237, Adjusted R-squared:  0.105
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.534  0.881  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      8.13e-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 32.92 1      5.738
## Year              32.92 16      1.115

```

Residuals from first author



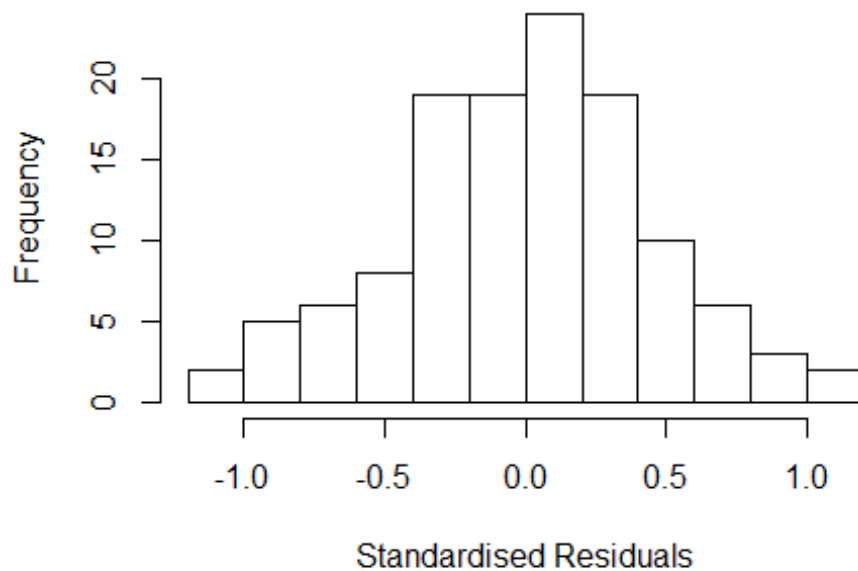
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.18e+00 -3.03e-01 4.44e-16 3.00e-01 1.01e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3885 0.1350 10.29 < 2e-16 ***
## FirstAuthorFemale1 0.1279 0.1253 1.02 0.310
## Year1997 -0.5386 0.2091 -2.58 0.011 *
## Year1998 -0.1186 0.1746 -0.68 0.498
## Year1999 -0.4309 0.3409 -1.26 0.209
## Year2000 -0.6505 0.1350 -4.82 4.9e-06 ***
## Year2001 -0.3894 0.1809 -2.15 0.034 *
## Year2002 -0.1769 0.3497 -0.51 0.614
## Year2003 -0.1591 0.1893 -0.84 0.403
## Year2004 -0.2118 0.2423 -0.87 0.384
## Year2005 0.0177 0.1961 0.09 0.928
## Year2006 -0.4377 0.1918 -2.28 0.025 *
```

```

## Year2007          -0.7656      0.1880    -4.07  9.0e-05 ***
## Year2008           0.1074      0.2445      0.44   0.661
## Year2009          -0.1424      0.1865     -0.76   0.447
## Year2010          -0.3949      0.2647     -1.49   0.139
## Year2011          -0.2346      0.2697     -0.87   0.386
## Year2012           0.1746      0.1484      1.18   0.242
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.472
## Multiple R-squared:  0.23,   Adjusted R-squared:  0.105
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 107 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.514  0.875   0.956   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      8.13e-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 2.102 1          1.450
## Year             2.102 16          1.023

```

Residuals from last author



```
## [1] "List of 0 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
## -1.16630 -0.29707  0.00589  0.28282  1.00889
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3775     0.1385   9.95 < 2e-16 ***
## LastAuthorFemale1  0.1687     0.1334   1.26  0.20894
## Year1997        -0.5143     0.2093  -2.46  0.01566 *
## Year1998        -0.0844     0.1706  -0.49  0.62178
## Year1999        -0.4297     0.3002  -1.43  0.15525
## Year2000        -0.6395     0.1385  -4.62  1.1e-05 ***
## Year2001        -0.3826     0.1810  -2.11  0.03689 *
## Year2002        -0.1647     0.3504  -0.47  0.63929
## Year2003        -0.1814     0.1982  -0.92  0.36215
## Year2004        -0.2112     0.2432  -0.87  0.38722
## Year2005         0.0449     0.1948   0.23  0.81825
## Year2006        -0.4303     0.1939  -2.22  0.02860 *
```

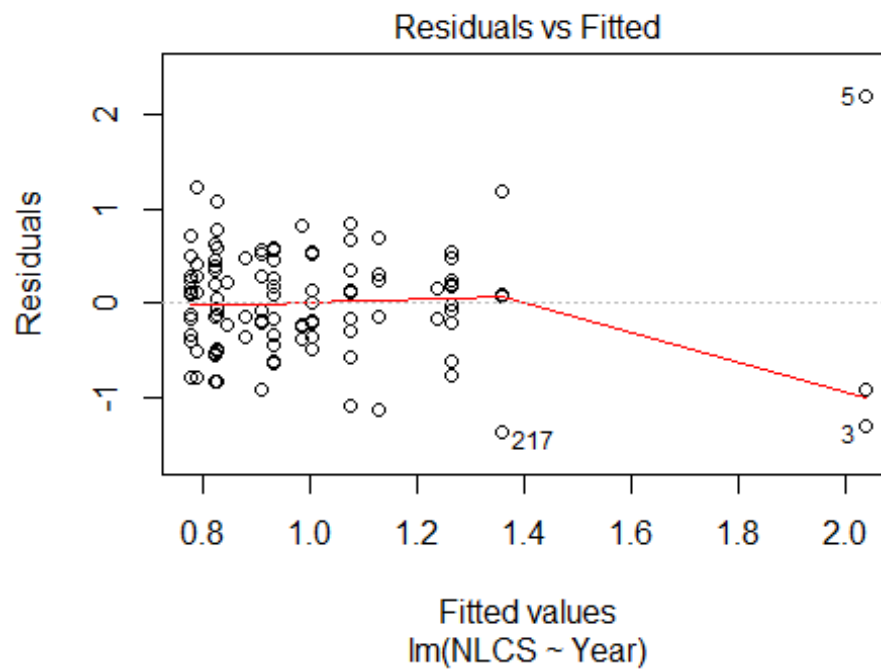


```

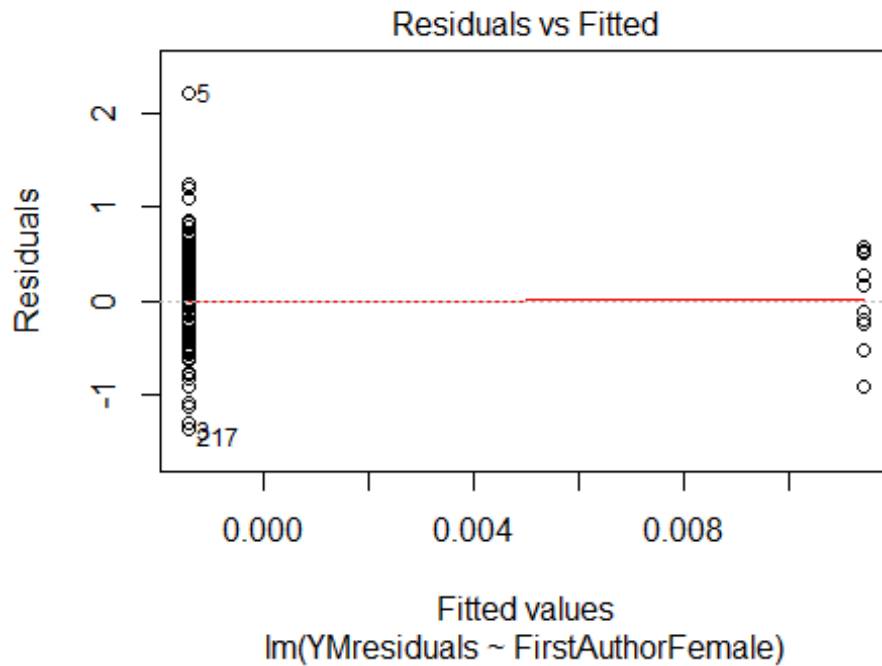
## Year2007          -0.7569      0.1925    -3.93   0.00015 ***
## Year2008           0.0657      0.2363      0.28   0.78166
## Year2009          -0.1203      0.1854     -0.65   0.51787
## Year2010          -0.4014      0.2806     -1.43   0.15561
## Year2011          -0.2274      0.2594     -0.88   0.38262
## Year2012           0.2495      0.1497      1.67   0.09856 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.236, Adjusted R-squared:  0.113
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.525  0.875   0.956   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      8.13e-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: 123"
## [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
##    5   10   10   15   17   17    7    6    9    5    4    8    8   12   12
## 2011 2012
##   13   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    8    3   10   10   11    6    3    7    4    2    5    2    9    4
## 2011 2012

```

```
##      7   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    8    3    8    9   11    6    3    6    4    2    3    2    8    4
## 2011 2012
##    5   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 = 24, df = 16, p-value = 0.08
```

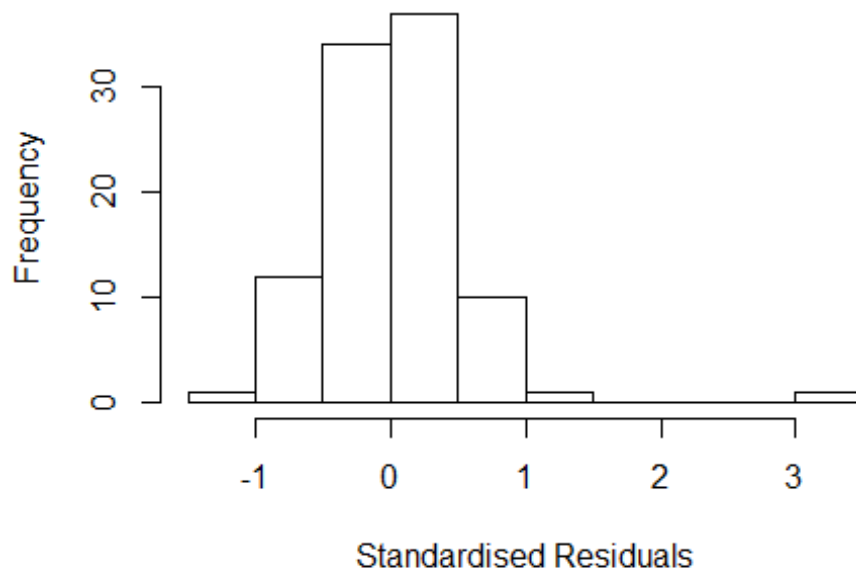


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.1, df = 1, p-value = 0.3
```



```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 1.73205080756888"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 52.88 1          7.272
## LastAuthorFemale  37.43 1          6.118
## UniqueAuthors    315.87 3          2.610
## Year              254.33 16          1.189
```

Residuals from first and last author and team size



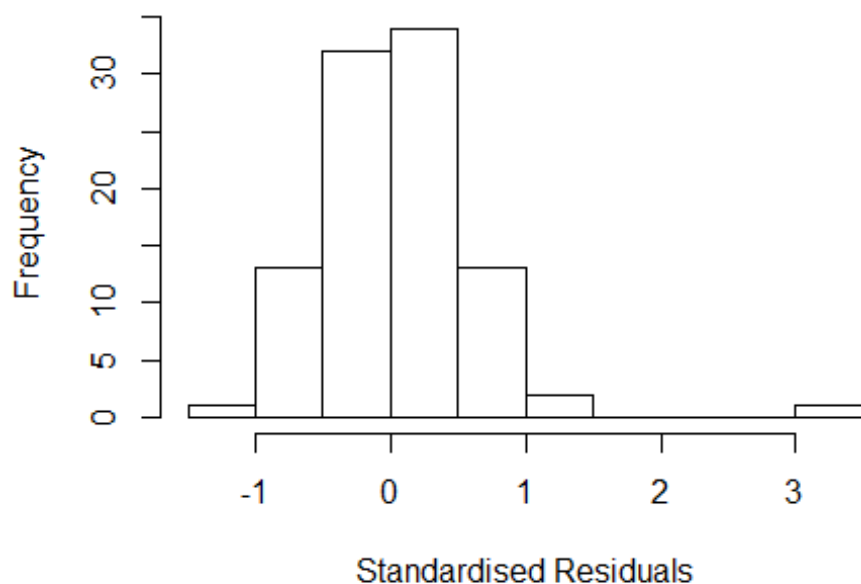
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 5 0030305457 4.238 1996      1804      3      3.146
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3038 -0.2755  0.0162  0.3115  3.1465
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.78446    0.07603   10.32 5.7e-16 ***
## FirstAuthorFemale1 -0.03127    0.16240   -0.19  0.848
## LastAuthorFemale1 -0.14850    0.13625   -1.09  0.279
## UniqueAuthors2     0.30708    0.14252    2.15  0.034 *
## UniqueAuthors3     0.82146    0.14828    5.54 4.4e-07 ***
## UniqueAuthors5     0.40505    0.23054    1.76  0.083 .
## Year1997          -0.00394    0.11175   -0.04  0.972
## Year1998           0.08499    0.23337    0.36  0.717
## Year1999           0.37486    0.19473    1.92  0.058 .
## Year2000          -0.17285    0.20490   -0.84  0.402
```

```

## Year2001      0.06217      0.16853      0.37      0.713
## Year2002     -0.32148      0.25345     -1.27      0.209
## Year2003      0.05247      0.25612      0.20      0.838
## Year2004     -0.30691      0.24551     -1.25      0.215
## Year2005      0.14366      0.26317      0.55      0.587
## Year2006      0.37575      0.18774      2.00      0.049 *
## Year2007     -0.05822      0.32523     -0.18      0.858
## Year2008     -0.07786      0.29728     -0.26      0.794
## Year2009     -0.16320      0.15121     -1.08      0.284
## Year2010      0.51937      0.72388      0.72      0.475
## Year2011      0.18026      0.11320      1.59      0.116
## Year2012     -0.01845      0.11231     -0.16      0.870
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared:  0.265, Adjusted R-squared:  0.056
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## observation 2 is an outlier with |weight| = 0 ( < 0.001);
## 9 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.481  0.905  0.958  0.927  0.988  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           1.04e-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 1.683 1           1.297
## LastAuthorFemale  1.811 1           1.346
## Year              2.645 16           1.031

```

Residuals from first and last author



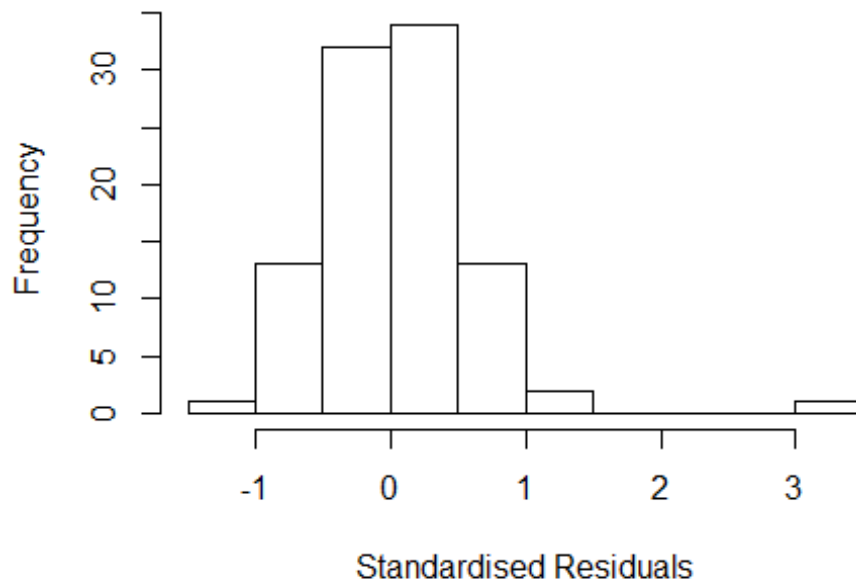
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 5 0030305457 4.238 1996      1804      3      3.3
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4647 -0.3319  0.0163  0.3771  3.3000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.93800    0.13787   6.80  2e-09 ***
## FirstAuthorFemale1  0.07859    0.14828   0.53   0.60
## LastAuthorFemale1  0.00406    0.17140   0.02   0.98
## Year1997         0.04888    0.19435   0.25   0.80
## Year1998        -0.06625    0.25719  -0.26   0.80
## Year1999         0.33307    0.22463   1.48   0.14
## Year2000        -0.15084    0.28696  -0.53   0.60
## Year2001        -0.00547    0.19858  -0.03   0.98
## Year2002        -0.09983    0.29812  -0.33   0.74
## Year2003        -0.10485    0.27587  -0.38   0.70
## Year2004        -0.35548    0.27449  -1.30   0.20
## Year2005         0.00445    0.28693   0.02   0.99
```

```

## Year2006          0.29947    0.19777    1.51    0.13
## Year2007          -0.11186    0.41648   -0.27    0.79
## Year2008          -0.13279    0.24999   -0.53    0.60
## Year2009           0.05073    0.23538    0.22    0.83
## Year2010           0.52674    0.59283    0.89    0.38
## Year2011           0.15603    0.21130    0.74    0.46
## Year2012          -0.08036    0.17829   -0.45    0.65
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.572
## Multiple R-squared:  0.132, Adjusted R-squared:  -0.0714
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 2 is an outlier with |weight| = 0 ( < 0.001);
## 8 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.492  0.915   0.959   0.934   0.988   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      1.04e-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 1.542  1          1.242
## Year              1.542 16          1.014

```

Residuals from first author



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 5 0030305457 4.238 1996      1804      3      3.3
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4697 -0.3320  0.0171  0.3759  3.3000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.93800    0.13795   6.80 1.9e-09 ***
## FirstAuthorFemale1 0.07960    0.14182   0.56  0.576
## Year1997        0.04899    0.19628   0.25  0.804
## Year1998       -0.06646    0.25752  -0.26  0.797
## Year1999        0.33327    0.22446   1.48  0.142
## Year2000       -0.15183    0.28779  -0.53  0.599
## Year2001       -0.00539    0.19889  -0.03  0.978
## Year2002       -0.09841    0.28883  -0.34  0.734
## Year2003       -0.10451    0.27628  -0.38  0.706
## Year2004       -0.35547    0.27543  -1.29  0.201
## Year2005        0.00317    0.28695   0.01  0.991
## Year2006        0.30150    0.17738   1.70  0.093 .
```

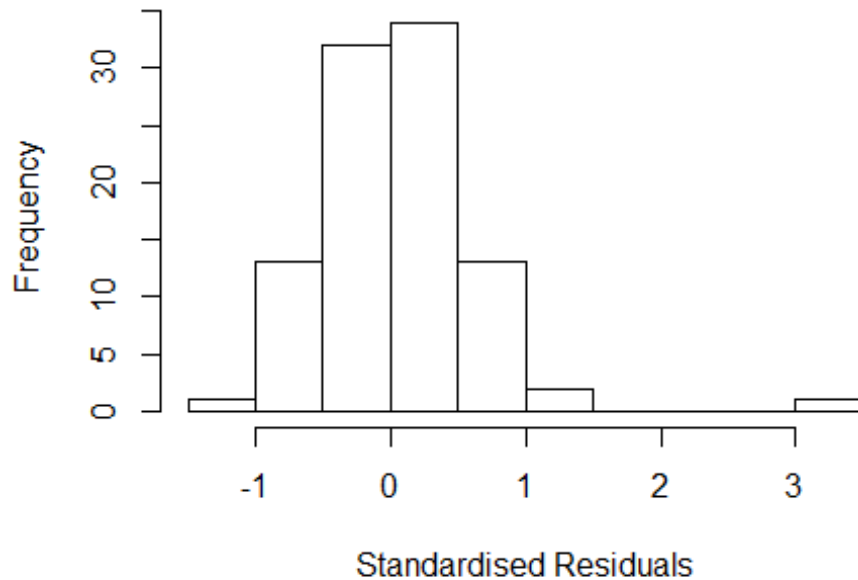


```

## Year2007      -0.11065    0.41815   -0.26    0.792
## Year2008      -0.13330    0.24959   -0.53    0.595
## Year2009       0.05127    0.23563    0.22    0.828
## Year2010       0.53169    0.59623    0.89    0.375
## Year2011       0.15649    0.20920    0.75    0.457
## Year2012      -0.07921    0.17749   -0.45    0.657
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.565
## Multiple R-squared:  0.133, Adjusted R-squared:  -0.0564
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 2 is an outlier with |weight| = 0 ( < 0.001);
## 8 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.478  0.913  0.958  0.932  0.988  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      1.04e-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 1.601 1          1.265
## Year              1.601 16          1.015

```

Residuals from last author



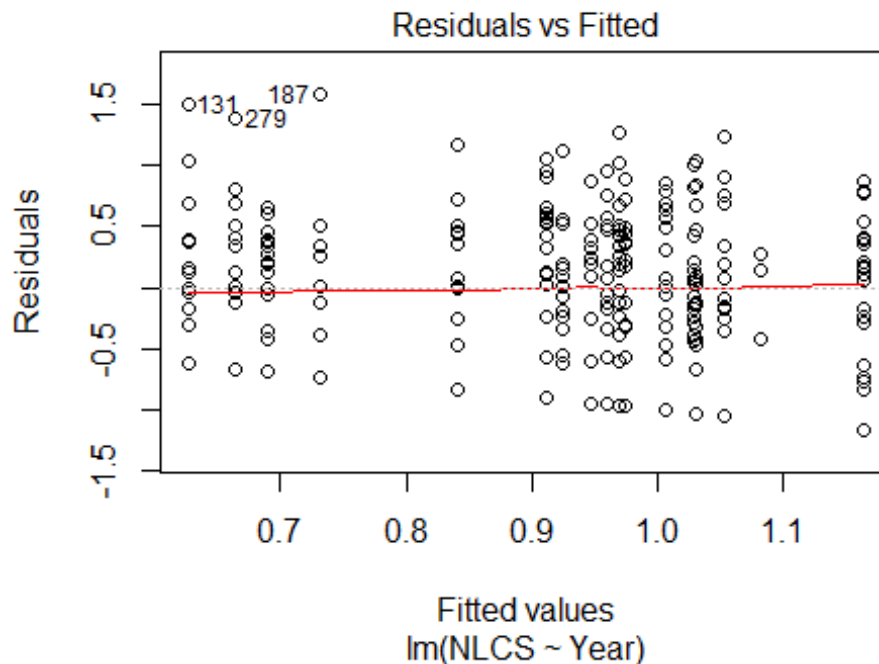
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 5 0030305457 4.238 1996      1804      3      3.3
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4714 -0.3338  0.0278  0.3570  3.3000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.93800    0.13798   6.80 1.9e-09 ***
## LastAuthorFemale1 0.02328    0.15940   0.15  0.88
## Year1997        0.05724    0.18994   0.30  0.76
## Year1998       -0.06653    0.25766  -0.26  0.80
## Year1999        0.35481    0.22195   1.60  0.11
## Year2000       -0.13607    0.28909  -0.47  0.64
## Year2001       -0.00536    0.19902  -0.03  0.98
## Year2002       -0.10319    0.29954  -0.34  0.73
## Year2003       -0.10439    0.27647  -0.38  0.71
## Year2004       -0.35547    0.27576  -1.29  0.20
## Year2005        0.00272    0.28715   0.01  0.99
## Year2006        0.28986    0.19955   1.45  0.15
```

```

## Year2007          -0.11023      0.41890    -0.26      0.79
## Year2008          -0.09350      0.21456    -0.44      0.66
## Year2009           0.05146      0.23575     0.22      0.83
## Year2010           0.53345      0.59749     0.89      0.37
## Year2011           0.18477      0.20607     0.90      0.37
## Year2012          -0.07177      0.17430    -0.41      0.68
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.562
## Multiple R-squared:  0.131, Adjusted R-squared:  -0.058
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 2 is an outlier with |weight| = 0 ( < 0.001);
## 8 weights are ~= 1. The remaining 87 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.473  0.913  0.951  0.931  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      1.04e-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: 96"
## [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
##   41   44   30   26   26   34   26   28   26   23   23   14   27   23   23
## 2011 2012
##   28   32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   15   26   17   16   10    3   19   25   18   15   19   14   21   17   18

```

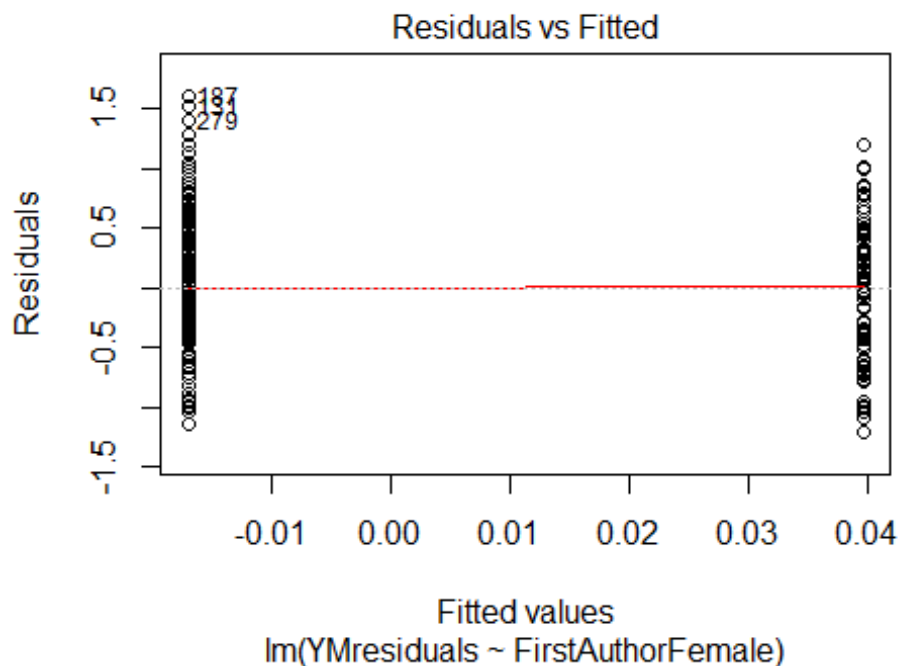
```
## 2011 2012
## 21 19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 25 17 14 9 3 18 24 18 12 18 14 21 16 18
## 2011 2012
## 20 19
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 12, df = 16, p-value = 0.8
```



```
##
## 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: 3.02241878101881"
## [1] "Male first author team size 2018 geometric mean: 2.44035291798841"
## 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 = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.82286471646455"
## [1] "Male last author team size 2018 geometric mean: 2.50892215513553"

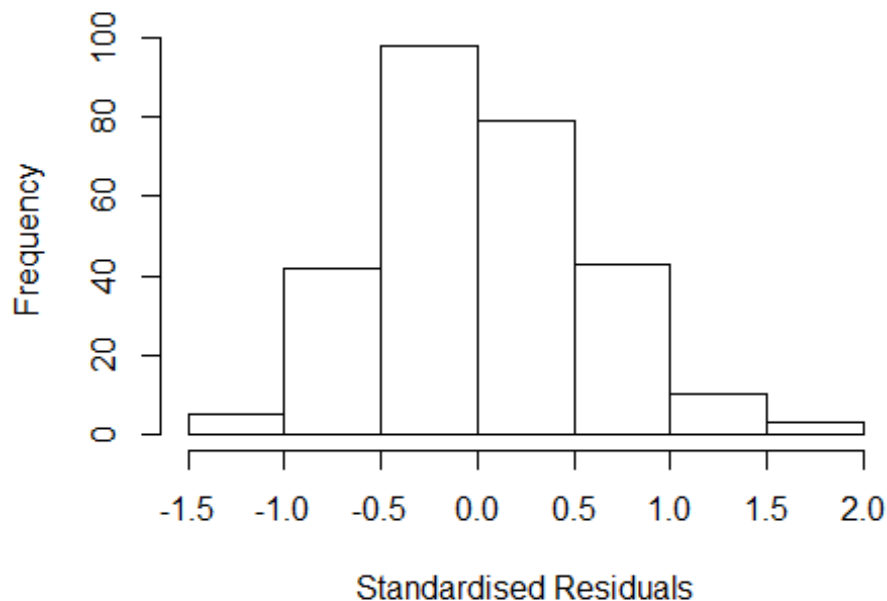
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 50, 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	1.900	1	1.379
LastAuthorFemale	1.738	1	1.318
UniqueAuthors	2.940	4	1.144
Year	5.882	16	1.057

Residuals from first and last author and team size



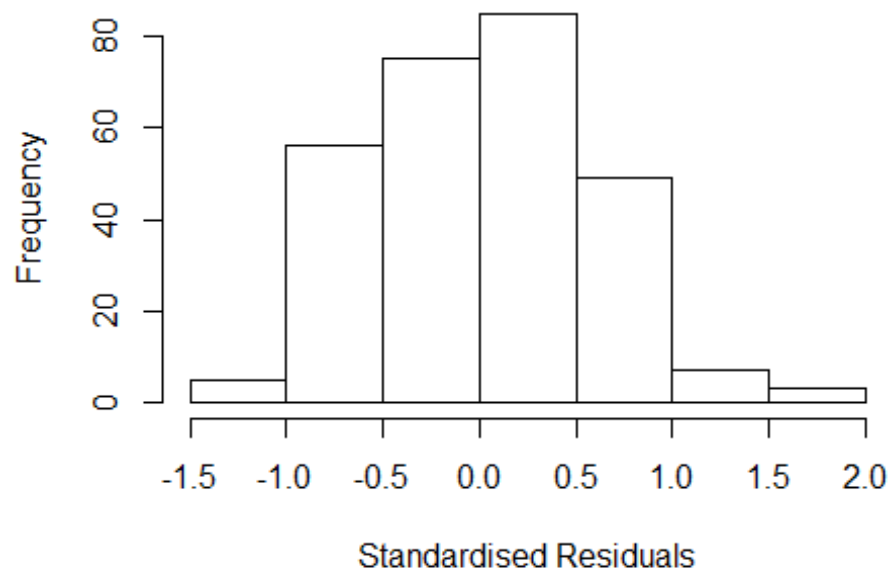
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0852 -0.3922 -0.0147 0.3709 1.8897
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0081 0.1676 6.02 6.1e-09 ***
## FirstAuthorFemale1 0.1554 0.0947 1.64 0.10185
## LastAuthorFemale1 -0.0820 0.0885 -0.93 0.35502
## UniqueAuthors2 0.2988 0.0834 3.58 0.00041 ***
## UniqueAuthors3 0.4157 0.0944 4.40 1.6e-05 ***
## UniqueAuthors4 0.3057 0.1328 2.30 0.02213 *
## UniqueAuthors5 0.5469 0.1725 3.17 0.00170 **
## Year1997 -0.0322 0.2254 -0.14 0.88653
## Year1998 -0.5661 0.2058 -2.75 0.00637 **
## Year1999 -0.3702 0.2577 -1.44 0.15217
```

```

## Year2000          -0.5847      0.2512    -2.33  0.02068 *
## Year2001          -0.2154      0.1824    -1.18  0.23867
## Year2002          -0.5602      0.2233    -2.51  0.01273 *
## Year2003          -0.3401      0.2324    -1.46  0.14460
## Year2004          -0.4529      0.2008    -2.26  0.02496 *
## Year2005          -0.2883      0.2364    -1.22  0.22369
## Year2006          -0.2217      0.2165    -1.02  0.30675
## Year2007          -0.2628      0.2346    -1.12  0.26376
## Year2008          -0.2622      0.2177    -1.20  0.22962
## Year2009          -0.2423      0.2015    -1.20  0.23018
## Year2010          -0.2172      0.1962    -1.11  0.26930
## Year2011          -0.2878      0.1967    -1.46  0.14465
## Year2012          -0.3079      0.2264    -1.36  0.17516
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.179, Adjusted R-squared:  0.108
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 254 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.217  0.875   0.945   0.909   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          3.57e-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 1.820 1          1.349
## LastAuthorFemale 1.688 1          1.299
## Year          1.809 16          1.019

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.2060 -0.4216 0.0194 0.4422 1.6739
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1286 0.1889 5.98 7.5e-09 ***
## FirstAuthorFemale1 0.1700 0.0996 1.71 0.089 .
## LastAuthorFemale1 -0.1216 0.0982 -1.24 0.216
## Year1997 0.0290 0.2358 0.12 0.902
## Year1998 -0.5712 0.2446 -2.34 0.020 *
## Year1999 -0.3444 0.2552 -1.35 0.178
## Year2000 -0.4895 0.3233 -1.51 0.131
## Year2001 -0.0996 0.2466 -0.40 0.687
## Year2002 -0.5895 0.2511 -2.35 0.020 *
## Year2003 -0.2806 0.2527 -1.11 0.268
## Year2004 -0.4266 0.2363 -1.81 0.072 .
## Year2005 -0.3373 0.2520 -1.34 0.182
```

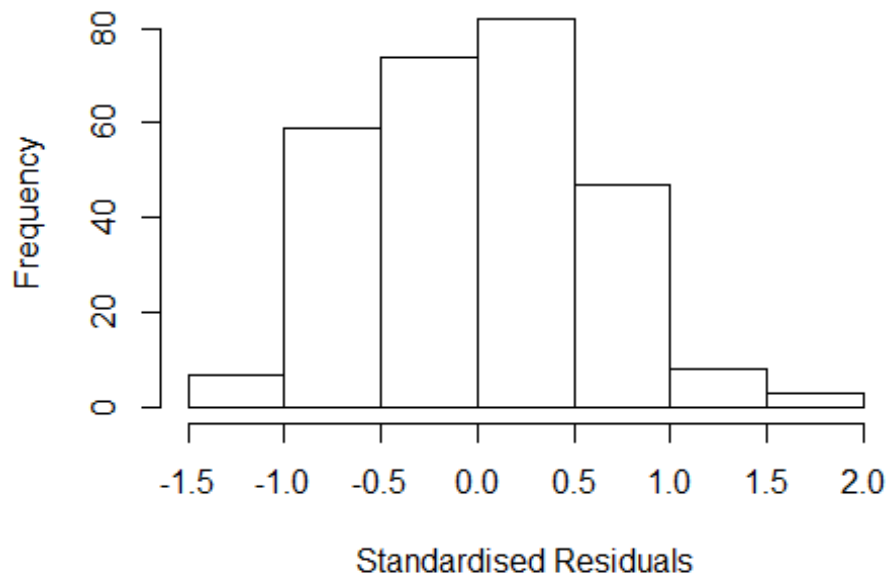


```

## Year2006          -0.1790      0.2366   -0.76    0.450
## Year2007          -0.1002      0.2508   -0.40    0.690
## Year2008          -0.2103      0.2425   -0.87    0.387
## Year2009          -0.1997      0.2214   -0.90    0.368
## Year2010          -0.1209      0.2143   -0.56    0.573
## Year2011          -0.1983      0.2281   -0.87    0.385
## Year2012          -0.1687      0.2421   -0.70    0.487
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.0914, Adjusted R-squared:  0.0287
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.421  0.880  0.942  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      3.57e-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 1.484 1      1.218
## Year              1.484 16      1.012

```

Residuals from first author



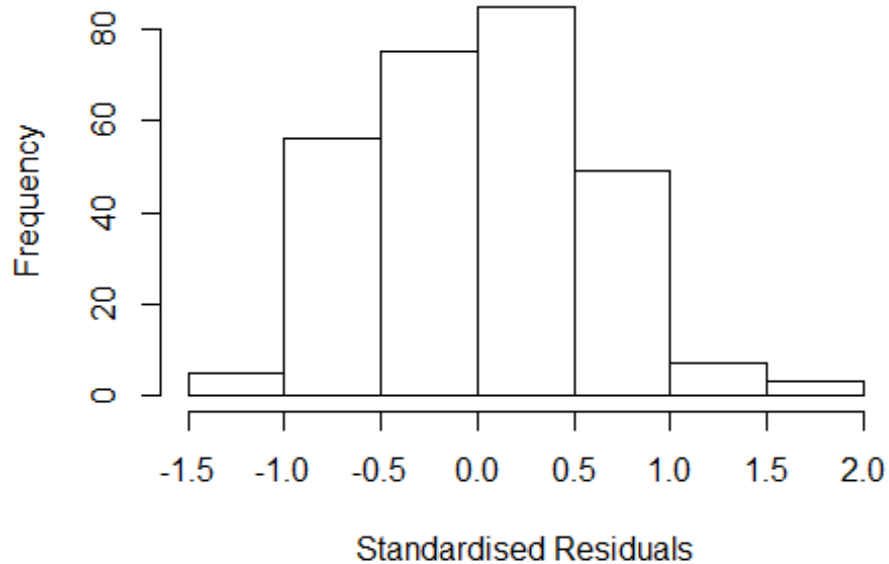
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.23578 -0.41894 0.00289 0.44832 1.70112
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1231 0.1906 5.89 1.2e-08 ***
## FirstAuthorFemale1 0.0974 0.0921 1.06 0.291
## Year1997 0.0153 0.2369 0.06 0.949
## Year1998 -0.5710 0.2466 -2.32 0.021 *
## Year1999 -0.3352 0.2591 -1.29 0.197
## Year2000 -0.5112 0.3197 -1.60 0.111
## Year2001 -0.0692 0.2549 -0.27 0.786
## Year2002 -0.5877 0.2555 -2.30 0.022 *
## Year2003 -0.2842 0.2509 -1.13 0.258
## Year2004 -0.4622 0.2302 -2.01 0.046 *
## Year2005 -0.3646 0.2496 -1.46 0.145
## Year2006 -0.1844 0.2382 -0.77 0.440
```

```

## Year2007          -0.1169      0.2494   -0.47    0.640
## Year2008          -0.2074      0.2450   -0.85    0.398
## Year2009          -0.2110      0.2218   -0.95    0.342
## Year2010          -0.1219      0.2147   -0.57    0.571
## Year2011          -0.2003      0.2284   -0.88    0.381
## Year2012          -0.1777      0.2443   -0.73    0.468
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.0877, Adjusted R-squared:  0.0285
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 262 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.408  0.882  0.947  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      3.57e-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 1.329 1          1.153
## Year            1.329 16          1.009

```

Residuals from last author



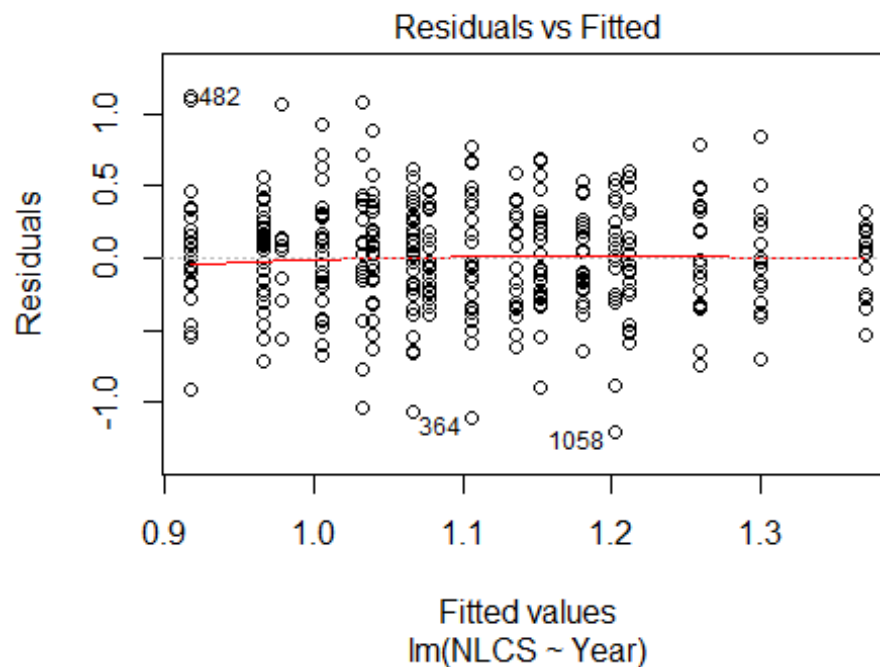
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.16841 -0.42270 0.00441 0.43327 1.66927
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1360 0.1864 6.09 3.9e-09 ***
## LastAuthorFemale1 -0.0178 0.0861 -0.21 0.836
## Year1997 0.0324 0.2360 0.14 0.891
## Year1998 -0.5622 0.2439 -2.30 0.022 *
## Year1999 -0.3157 0.2573 -1.23 0.221
## Year2000 -0.4923 0.3101 -1.59 0.114
## Year2001 -0.0491 0.2586 -0.19 0.849
## Year2002 -0.5419 0.2460 -2.20 0.028 *
## Year2003 -0.2569 0.2441 -1.05 0.294
## Year2004 -0.4257 0.2314 -1.84 0.067 .
## Year2005 -0.3011 0.2475 -1.22 0.225
## Year2006 -0.1668 0.2315 -0.72 0.472
```

```

## Year2007          -0.1110      0.2500   -0.44    0.658
## Year2008          -0.1682      0.2382   -0.71    0.481
## Year2009          -0.1838      0.2169   -0.85    0.398
## Year2010          -0.1219      0.2092   -0.58    0.561
## Year2011          -0.2066      0.2262   -0.91    0.362
## Year2012          -0.1459      0.2392   -0.61    0.542
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.613
## Multiple R-squared:  0.0807, Adjusted R-squared:  0.021
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 260 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.438  0.888  0.952  0.921  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      3.57e-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: 280"
## [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
##   45   49   42   34   51   62   34   43   52   43   43   44   42   23   45
## 2011 2012
##   42   45
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   20   29   24   18   27   25   20   27   32   18   31   22   24   10   16
## 2011 2012

```

```
## 16 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 16 25 23 17 25 25 15 25 27 16 26 19 22 9 16
## 2011 2012
## 13 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 = 23, df = 16, p-value = 0.1
```



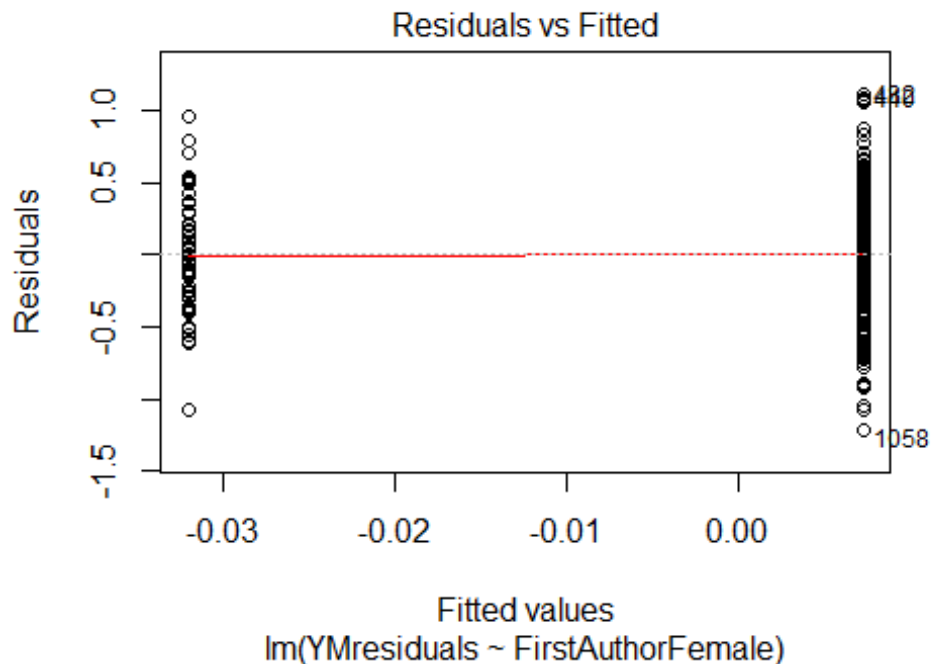
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.031, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 3.11337964201891"
## [1] "Male first author team size 2018 geometric mean: 2.10377803975693"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

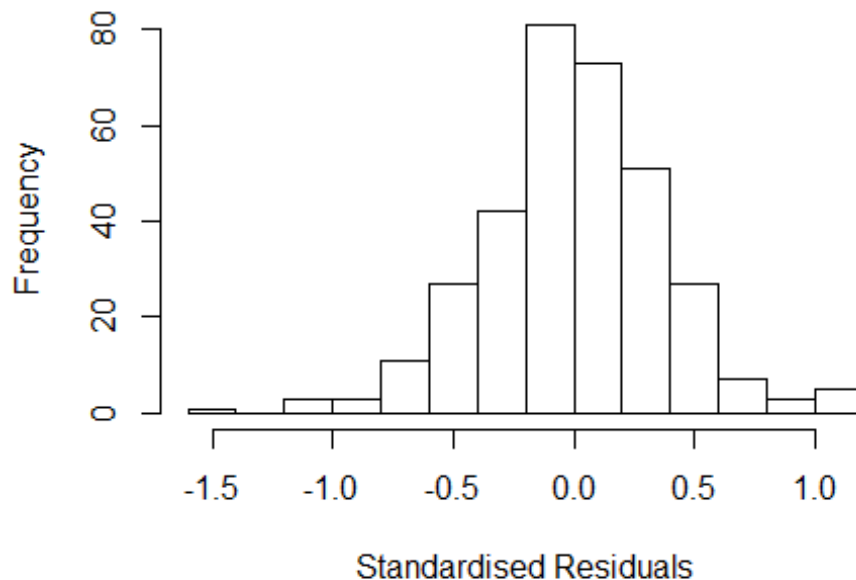
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male last author team size 2018 geometric mean: 2.5137684498864"

## 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 = 51, 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 1.345  1      1.160
## LastAuthorFemale  1.140  1      1.068
## UniqueAuthors    2.575  4      1.126
## Year              3.610 16      1.041
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.4151 -0.2074 -0.0011  0.2266  1.1592
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3320    0.1177   11.32  <2e-16 ***
## FirstAuthorFemale1 -0.0393    0.0575   -0.68  0.4954
## LastAuthorFemale1  -0.0047    0.0565   -0.08  0.9337
## UniqueAuthors2    -0.0591    0.0724   -0.82  0.4152
## UniqueAuthors3     0.0424    0.0717    0.59  0.5552
## UniqueAuthors4    -0.1061    0.0701   -1.51  0.1311
## UniqueAuthors5    -0.0578    0.0727   -0.79  0.4275
## Year1997          -0.3412    0.1383   -2.47  0.0142 *
## Year1998          -0.1353    0.1370   -0.99  0.3242
## Year1999          -0.2950    0.1692   -1.74  0.0822 .
```

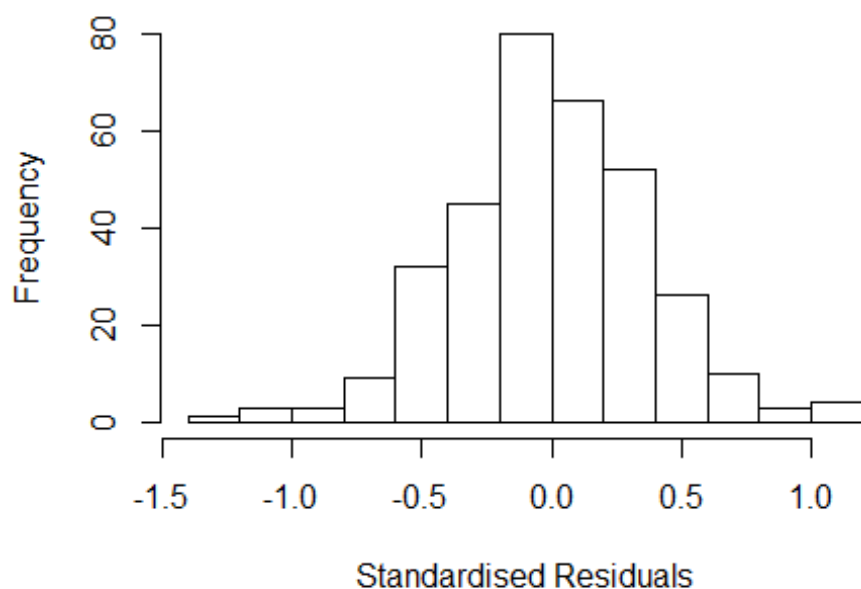


```

## Year2000          -0.2438      0.1272    -1.92    0.0562 .
## Year2001          -0.1728      0.1374    -1.26    0.2096
## Year2002           0.0763      0.1587     0.48    0.6310
## Year2003          -0.4242      0.1265    -3.35    0.0009 ***
## Year2004          -0.3136      0.1164    -2.69    0.0074 **
## Year2005          -0.1449      0.1467    -0.99    0.3240
## Year2006          -0.2122      0.1261    -1.68    0.0935 .
## Year2007          -0.2374      0.1191    -1.99    0.0470 *
## Year2008          -0.1369      0.1198    -1.14    0.2542
## Year2009          -0.3580      0.1585    -2.26    0.0246 *
## Year2010           0.0105      0.1389     0.08    0.9396
## Year2011           0.0637      0.1187     0.54    0.5917
## Year2012           0.0830      0.1350     0.62    0.5390
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.348
## Multiple R-squared:  0.149, Adjusted R-squared:  0.0884
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 312 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0607 0.8580 0.9510 0.8950 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          2.99e-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 1.277 1          1.130
## LastAuthorFemale  1.134 1          1.065
## Year              1.440 16          1.011

```

Residuals from first and last author



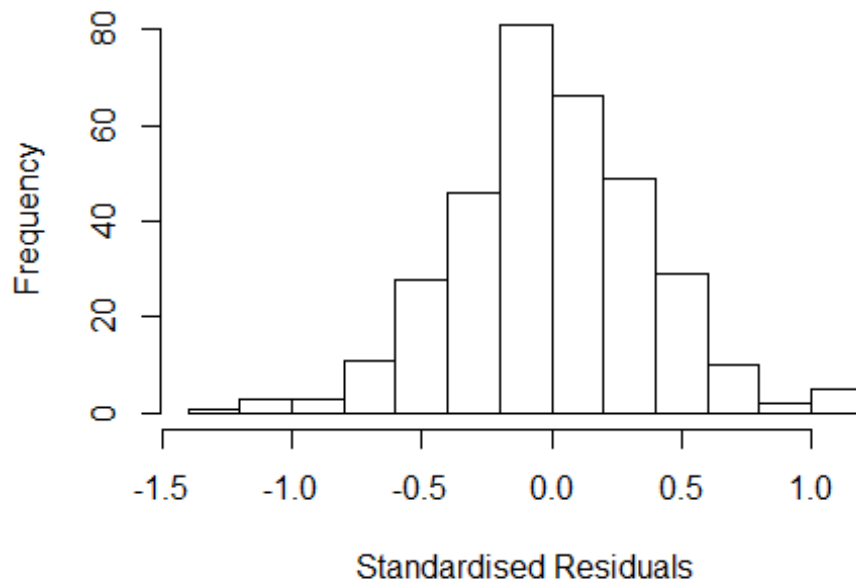
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3521 -0.2222 -0.0145 0.2355 1.1405
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.28297 0.09942 12.90 <2e-16 ***
## FirstAuthorFemale1 -0.04018 0.05650 -0.71 0.4775
## LastAuthorFemale1 -0.01418 0.05713 -0.25 0.8041
## Year1997 -0.30802 0.13278 -2.32 0.0210 *
## Year1998 -0.10845 0.13116 -0.83 0.4090
## Year1999 -0.26738 0.16473 -1.62 0.1056
## Year2000 -0.22994 0.12676 -1.81 0.0706 .
## Year2001 -0.15672 0.13790 -1.14 0.2566
## Year2002 0.07833 0.15535 0.50 0.6145
## Year2003 -0.37792 0.12470 -3.03 0.0026 **
## Year2004 -0.29031 0.11415 -2.54 0.0115 *
## Year2005 -0.13891 0.14530 -0.96 0.3398
```

```

## Year2006          -0.20641      0.12598      -1.64      0.1023
## Year2007          -0.22436      0.11581      -1.94      0.0536 .
## Year2008          -0.12539      0.11503      -1.09      0.2765
## Year2009          -0.38250      0.16312      -2.34      0.0197 *
## Year2010           0.00856      0.13566       0.06      0.9497
## Year2011           0.08844      0.11661       0.76      0.4488
## Year2012           0.06914      0.12927       0.53      0.5931
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0779
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 311 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.137  0.875   0.953   0.902   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      2.99e-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 1.268 1      1.126
## Year              1.268 16      1.007

```

Residuals from first author



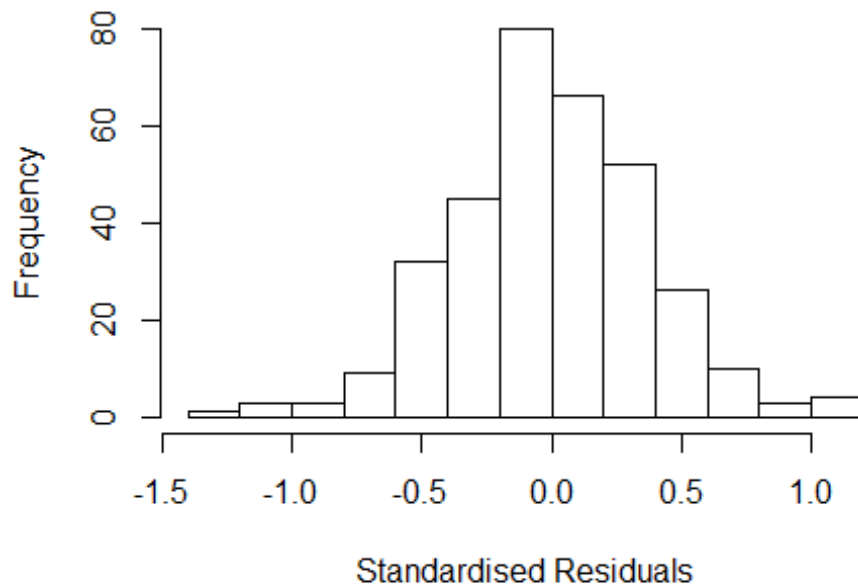
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3510 -0.2221 -0.0115 0.2347 1.1423
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.28299 0.09992 12.84 <2e-16 ***
## FirstAuthorFemale1 -0.04275 0.05695 -0.75 0.4533
## Year1997 -0.31045 0.13205 -2.35 0.0193 *
## Year1998 -0.11026 0.13070 -0.84 0.3995
## Year1999 -0.26945 0.16353 -1.65 0.1004
## Year2000 -0.23317 0.12496 -1.87 0.0630 .
## Year2001 -0.15894 0.13686 -1.16 0.2464
## Year2002 0.07624 0.15459 0.49 0.6222
## Year2003 -0.37957 0.12529 -3.03 0.0027 **
## Year2004 -0.29153 0.11368 -2.56 0.0108 *
## Year2005 -0.13986 0.14511 -0.96 0.3359
## Year2006 -0.20805 0.12514 -1.66 0.0974 .
```

```

## Year2007          -0.22655    0.11545   -1.96    0.0506 .
## Year2008          -0.12608    0.11501   -1.10    0.2738
## Year2009          -0.38430    0.16333   -2.35    0.0192 *
## Year2010           0.00737    0.13646    0.05    0.9570
## Year2011           0.08881    0.11688    0.76    0.4479
## Year2012           0.06802    0.12762    0.53    0.5944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.362
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0813
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 312 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.134  0.877   0.953   0.902   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      2.99e-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 1.121 1          1.059
## Year            1.121 16          1.004

```

Residuals from last author



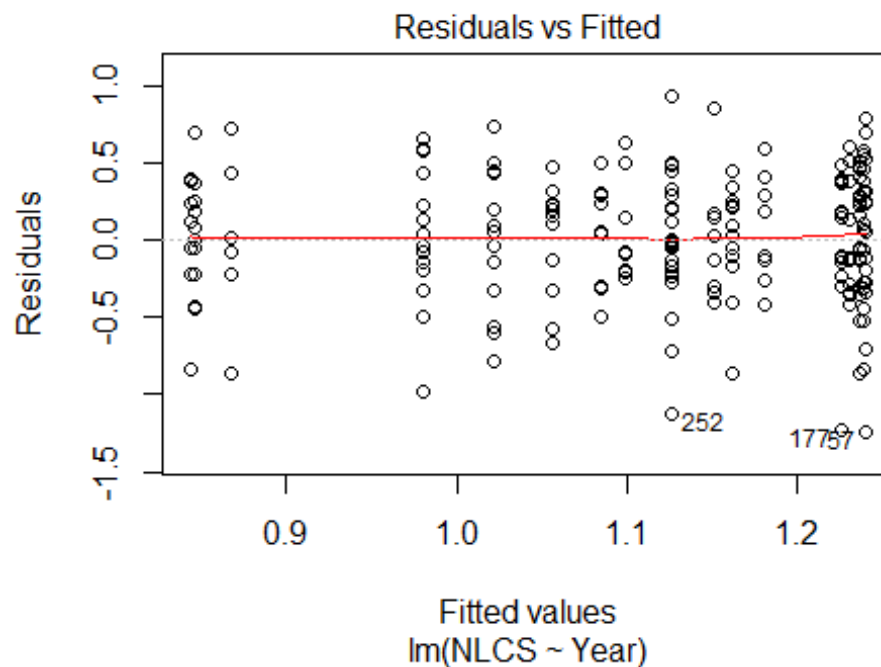
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3397 -0.2355 -0.0302  0.2381  1.1506
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2768    0.0982   13.00  <2e-16 ***
## LastAuthorFemale1 -0.0235    0.0586   -0.40  0.6892
## Year1997         -0.3032    0.1317   -2.30  0.0219 *
## Year1998         -0.1076    0.1307   -0.82  0.4107
## Year1999         -0.2640    0.1646   -1.60  0.1097
## Year2000         -0.2297    0.1270   -1.81  0.0715 .
## Year2001         -0.1537    0.1368   -1.12  0.2620
## Year2002          0.0759    0.1558    0.49  0.6264
## Year2003         -0.3782    0.1247   -3.03  0.0026 **
## Year2004         -0.2917    0.1144   -2.55  0.0113 *
## Year2005         -0.1352    0.1454   -0.93  0.3531
## Year2006         -0.2048    0.1259   -1.63  0.1046
```

```

## Year2007          -0.2267      0.1164    -1.95    0.0522 .
## Year2008          -0.1297      0.1153    -1.13    0.2613
## Year2009          -0.3864      0.1614    -2.39    0.0172 *
## Year2010           0.0119      0.1347     0.09    0.9294
## Year2011           0.0893      0.1176     0.76    0.4482
## Year2012           0.0629      0.1290     0.49    0.6260
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.366
## Multiple R-squared:  0.126, Adjusted R-squared:  0.0785
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 316 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.153  0.877  0.957  0.905  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      2.99e-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: 334"
## [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
##   21   24   26   30   27   28   30   19   19   18   20   21   14   18   17
## 2011 2012
##   24   18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13   13   10   19   13   10   22   8   12   11   13   10   7   11   6
## 2011 2012

```

```
## 17 10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 12 9 16 13 10 22 7 11 9 9 10 7 11 6
## 2011 2012
## 17 7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 7.8, df = 16, p-value = 1
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.79, df = 1, p-value = 0.4

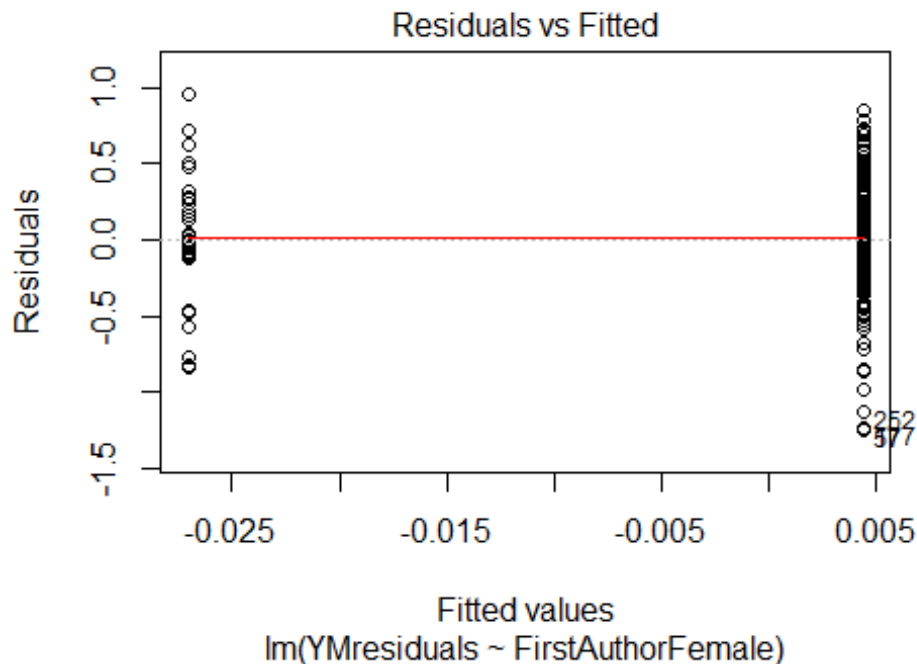
## [1] "Female first author team size 2018 geometric mean: 4.89897948556636"
## [1] "Male first author team size 2018 geometric mean: 2.62529810306986"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



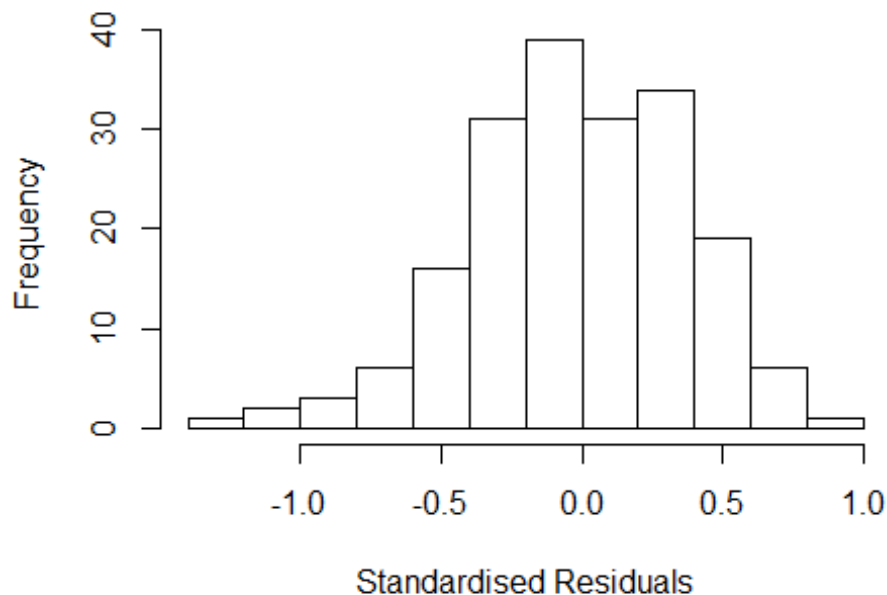
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 2.90516341745842"

## 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 = 6, 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 1.482  1      1.217
## LastAuthorFemale  1.626  1      1.275
## UniqueAuthors    5.030  4      1.224
## Year              7.750 16      1.066
```

Residuals from first and last author and team size



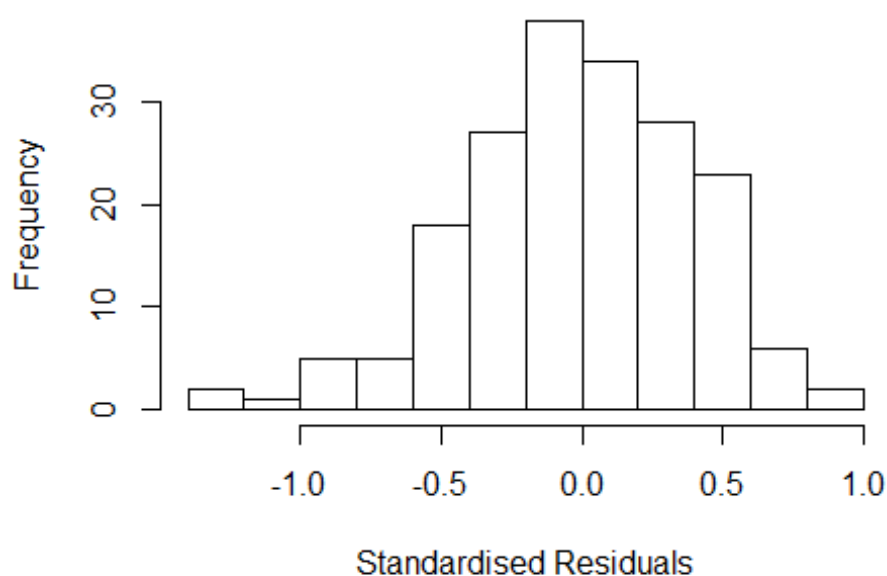
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2671 -0.2755 -0.0213 0.3064 0.8397
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.20177 0.16491 7.29 1.2e-11 ***
## FirstAuthorFemale1 -0.00798 0.09606 -0.08 0.9339
## LastAuthorFemale1 -0.06252 0.10315 -0.61 0.5453
## UniqueAuthors2 0.08383 0.09887 0.85 0.3977
## UniqueAuthors3 -0.09094 0.09781 -0.93 0.3538
## UniqueAuthors4 0.03404 0.10989 0.31 0.7571
## UniqueAuthors5 0.32155 0.12142 2.65 0.0089 **
## Year1997 0.06151 0.20706 0.30 0.7668
## Year1998 0.06689 0.19311 0.35 0.7295
## Year1999 0.04474 0.17034 0.26 0.7931
```

```

## Year2000      0.06537    0.17866    0.37    0.7149
## Year2001     -0.05709    0.18661   -0.31    0.7600
## Year2002     -0.10128    0.15887   -0.64    0.5247
## Year2003     -0.11729    0.22513   -0.52    0.6031
## Year2004     -0.07331    0.17570   -0.42    0.6770
## Year2005     -0.22817    0.20263   -1.13    0.2618
## Year2006     -0.33831    0.19627   -1.72    0.0866 .
## Year2007     -0.12337    0.17006   -0.73    0.4692
## Year2008     -0.41173    0.19912   -2.07    0.0402 *
## Year2009     -0.38568    0.18301   -2.11    0.0366 *
## Year2010     -0.35019    0.27298   -1.28    0.2013
## Year2011     -0.30357    0.16789   -1.81    0.0724 .
## Year2012     -0.12771    0.19300   -0.66    0.5091
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.165, Adjusted R-squared:  0.0538
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 172 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.311  0.897  0.945  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      5.29e-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 1.551 1      1.246
## LastAuthorFemale  1.469 1      1.212
## Year              2.214 16      1.025

```

Residuals from first and last author



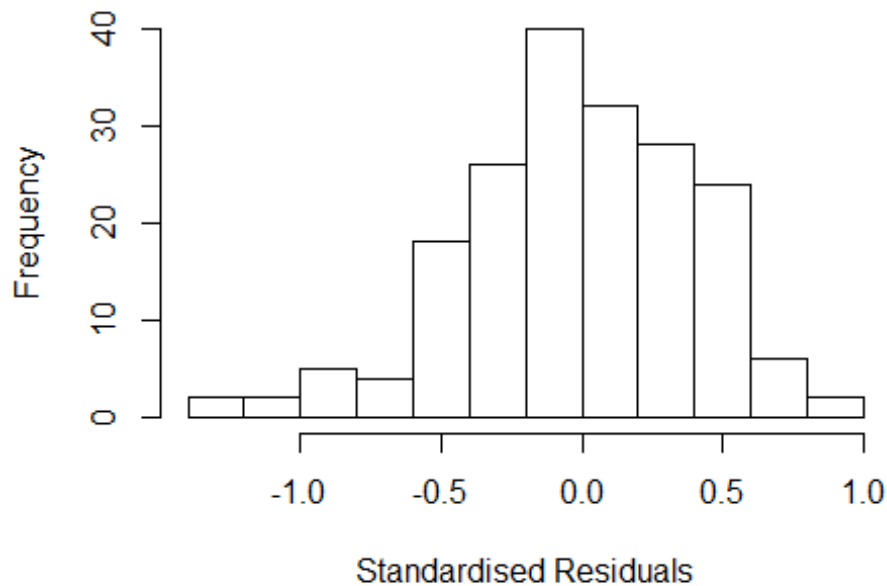
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.29123 -0.26664 -0.00305 0.29736 0.99252
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.25464 0.13792 9.10 2.4e-16 ***
## FirstAuthorFemale1 0.01481 0.10344 0.14 0.886
## LastAuthorFemale1 -0.10348 0.10380 -1.00 0.320
## Year1997 -0.02075 0.20142 -0.10 0.918
## Year1998 0.00128 0.20120 0.01 0.995
## Year1999 0.03326 0.16960 0.20 0.845
## Year2000 0.03659 0.16561 0.22 0.825
## Year2001 -0.07147 0.17963 -0.40 0.691
## Year2002 -0.10349 0.16169 -0.64 0.523
## Year2003 -0.15352 0.20788 -0.74 0.461
## Year2004 -0.07997 0.17335 -0.46 0.645
## Year2005 -0.23126 0.19687 -1.17 0.242
```

```

## Year2006      -0.26858    0.21785   -1.23    0.219
## Year2007      -0.15142    0.17125   -0.88    0.378
## Year2008      -0.36799    0.19661   -1.87    0.063 .
## Year2009      -0.39741    0.18405   -2.16    0.032 *
## Year2010      -0.34121    0.25888   -1.32    0.189
## Year2011      -0.25628    0.18075   -1.42    0.158
## Year2012      -0.09496    0.18759   -0.51    0.613
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.11, Adjusted R-squared:  0.0156
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 181 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.309  0.897  0.954  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      5.29e-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 1.546 1      1.243
## Year      1.546 16      1.014

```

Residuals from first author



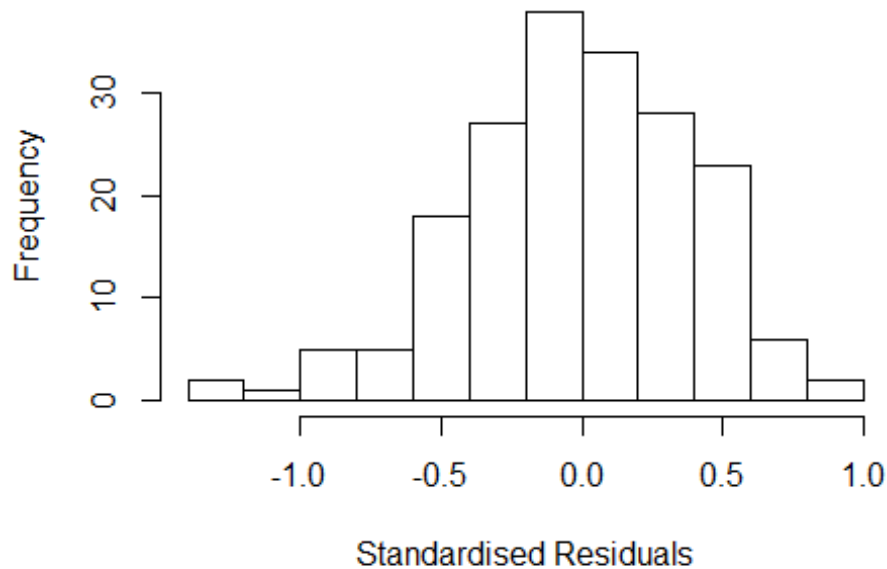
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2894 -0.2645 -0.0205 0.2783 0.9316
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.25405 0.13724 9.14 <2e-16 ***
## FirstAuthorFemale1 -0.02808 0.10693 -0.26 0.793
## Year1997 -0.02053 0.20245 -0.10 0.919
## Year1998 0.00458 0.19673 0.02 0.981
## Year1999 0.02498 0.16877 0.15 0.882
## Year2000 0.03537 0.16541 0.21 0.831
## Year2001 -0.07781 0.17880 -0.44 0.664
## Year2002 -0.10256 0.16174 -0.63 0.527
## Year2003 -0.14402 0.20968 -0.69 0.493
## Year2004 -0.08273 0.17333 -0.48 0.634
## Year2005 -0.23839 0.19391 -1.23 0.221
## Year2006 -0.28795 0.21186 -1.36 0.176
```

```

## Year2007      -0.17328    0.16937   -1.02    0.308
## Year2008      -0.36948    0.19928   -1.85    0.065 .
## Year2009      -0.40917    0.18057   -2.27    0.025 *
## Year2010      -0.35707    0.26280   -1.36    0.176
## Year2011      -0.24907    0.17967   -1.39    0.167
## Year2012      -0.10974    0.18769   -0.58    0.560
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.105, Adjusted R-squared:  0.0157
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 178 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.326  0.895   0.956   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      5.29e-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 1.416 1          1.190
## Year            1.416 16          1.011

```

Residuals from last author



```
## [1] "List of 0 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
## -1.28949 -0.26585 -0.00706 0.29793 0.99699
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.25407 0.13724 9.14 <2e-16 ***
## LastAuthorFemale1 -0.09549 0.10947 -0.87 0.384
## Year1997 -0.02152 0.20125 -0.11 0.915
## Year1998 0.00474 0.19910 0.02 0.981
## Year1999 0.03248 0.16896 0.19 0.848
## Year2000 0.03542 0.16544 0.21 0.831
## Year2001 -0.06989 0.17832 -0.39 0.696
## Year2002 -0.10057 0.16084 -0.63 0.533
## Year2003 -0.14892 0.20644 -0.72 0.472
## Year2004 -0.07802 0.17139 -0.46 0.650
## Year2005 -0.23039 0.19611 -1.17 0.242
## Year2006 -0.26880 0.21567 -1.25 0.214
```



```

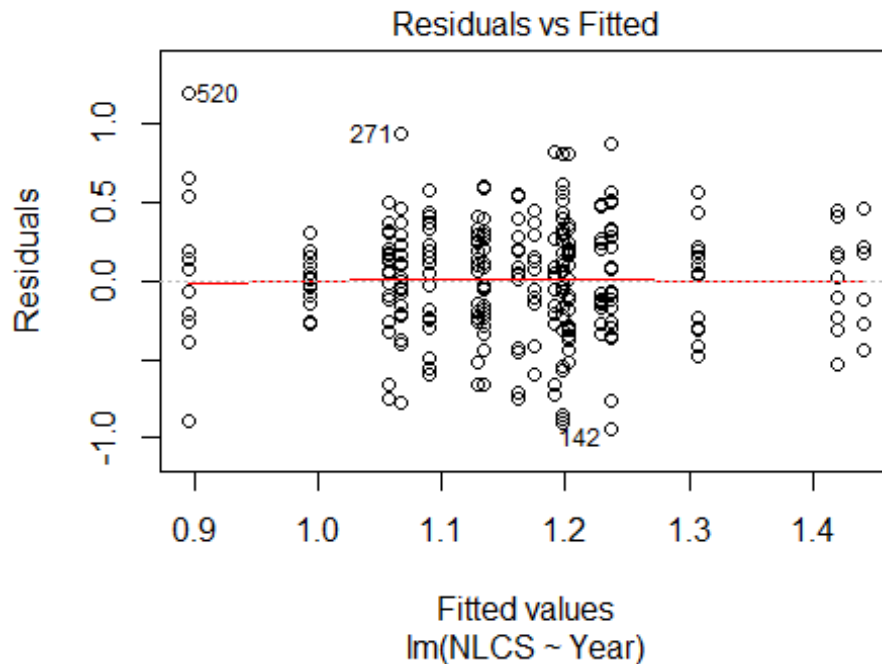
## Year2007          -0.15214      0.17044    -0.89      0.373
## Year2008          -0.36822      0.19592    -1.88      0.062 .
## Year2009          -0.39528      0.18227    -2.17      0.031 *
## Year2010          -0.33940      0.25698    -1.32      0.188
## Year2011          -0.25160      0.17265    -1.46      0.147
## Year2012          -0.09282      0.18629    -0.50      0.619
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.109, Adjusted R-squared:  0.0204
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 179 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.325  0.898  0.955  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      5.29e-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: 189"
## [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 2011
##    3    2    1   10    3    6    1    1    4    2    3    1    3    4    2
## 2012
##    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011
##    2    2    0    5    2    2    1    1    2    2    2    1    3    4    2
## 2012

```

```

##      5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011
##      2      2      0      2      2      2      1      1      2      2      1      1      2      4      2
## 2012
##      4
## [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: 2.05976714390712"
## [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: 30"
## [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
##      27      25      18      36      17      28      25      26      20      27      21      40      36      25      21
## 2011 2012
##      49      37
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      14      13      9      21      6      12      12      20      14      22      15      21      20      15      13
## 2011 2012
##      28      29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      14      11      9      18      6      11      12      20      14      18      10      20      20      14      13
## 2011 2012
##      25      25
## [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

```



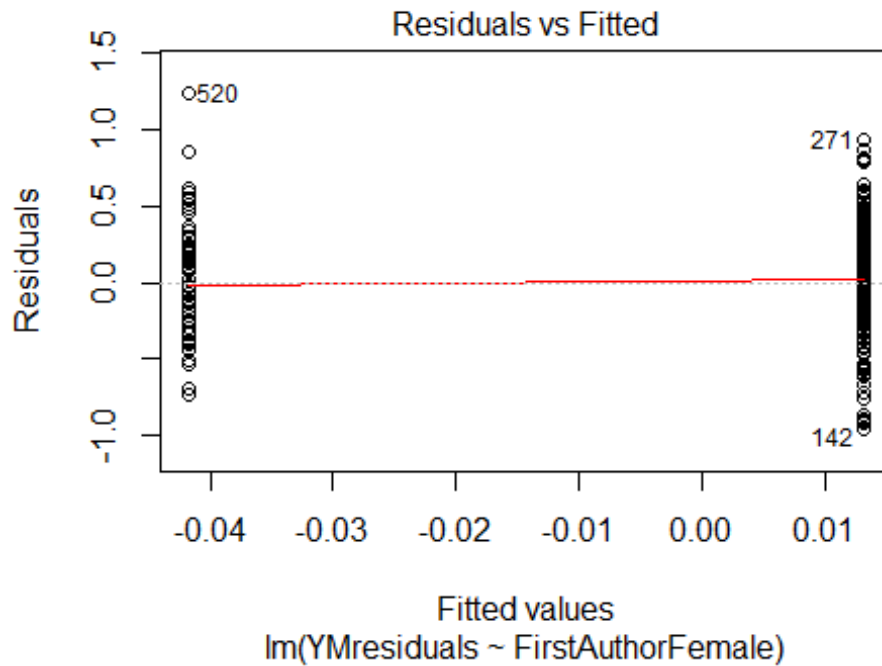
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0021, df = 1, p-value = 1

## [1] "Female first author team size 2018 geometric mean: 2.37956557896877"
## [1] "Male first author team size 2018 geometric mean: 3.03076530941501"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

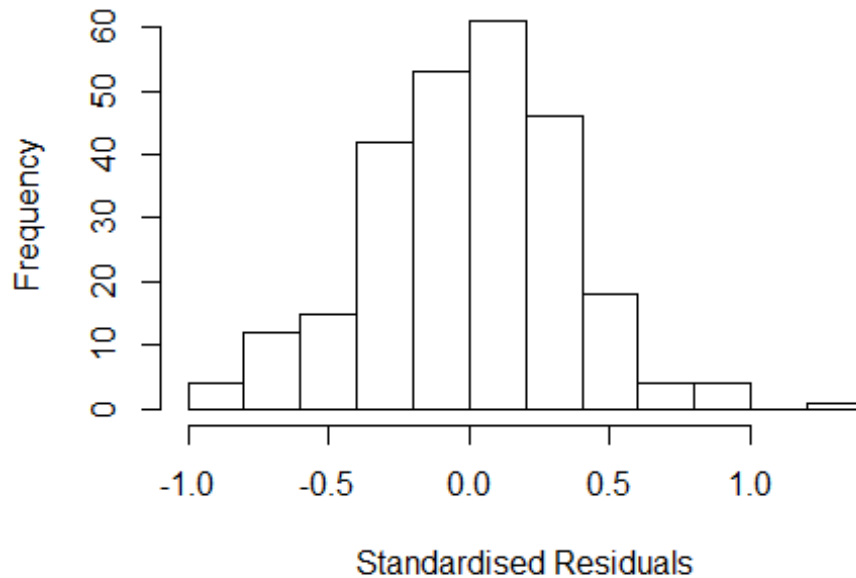
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 60, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.77274593888539"
## [1] "Male last author team size 2018 geometric mean: 2.89294274247302"

## 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 = 76, 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 1.967 1          1.402
## LastAuthorFemale  1.768 1          1.330
## UniqueAuthors     5.140 4          1.227
## Year              8.813 16          1.070
```

Residuals from first and last author and team size



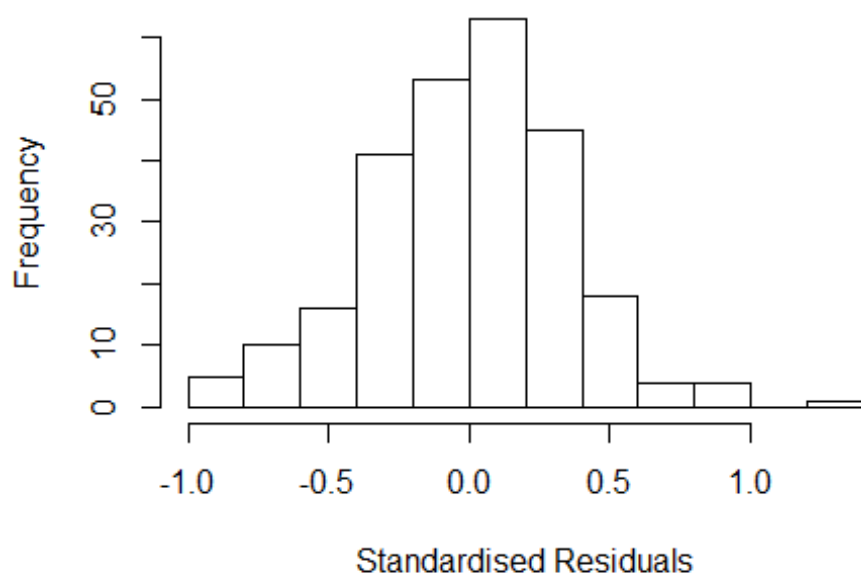
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9970 -0.2281 0.0108 0.2118 1.2048
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21325 0.13758 8.82 2.6e-16 ***
## FirstAuthorFemale1 -0.17383 0.06626 -2.62 0.0093 **
## LastAuthorFemale1 0.14503 0.05811 2.50 0.0132 *
## UniqueAuthors2 -0.01449 0.06794 -0.21 0.8313
## UniqueAuthors3 0.03380 0.06726 0.50 0.6158
## UniqueAuthors4 -0.08100 0.08524 -0.95 0.3429
## UniqueAuthors5 0.00392 0.07541 0.05 0.9586
## Year1997 -0.21786 0.14376 -1.52 0.1310
## Year1998 0.25353 0.17862 1.42 0.1571
## Year1999 0.04427 0.17761 0.25 0.8034
```

```

## Year2000      0.24027    0.19533    1.23    0.2199
## Year2001     -0.06570    0.16356   -0.40    0.6883
## Year2002     -0.09820    0.19888   -0.49    0.6219
## Year2003     -0.15316    0.15495   -0.99    0.3239
## Year2004     -0.00703    0.15602   -0.05    0.9641
## Year2005     -0.05778    0.15788   -0.37    0.7147
## Year2006     -0.02641    0.15765   -0.17    0.8671
## Year2007     -0.09546    0.15587   -0.61    0.5408
## Year2008     -0.11975    0.15216   -0.79    0.4321
## Year2009     -0.33603    0.21343   -1.57    0.1167
## Year2010      0.15126    0.15928    0.95    0.3433
## Year2011      0.09564    0.15248    0.63    0.5311
## Year2012      0.07878    0.15985    0.49    0.6226
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.348
## Multiple R-squared:  0.148, Adjusted R-squared:  0.0694
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 241 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.206  0.890  0.957  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      3.85e-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 1.893 1      1.376
## LastAuthorFemale  1.707 1      1.306
## Year              2.036 16      1.022

```

Residuals from first and last author



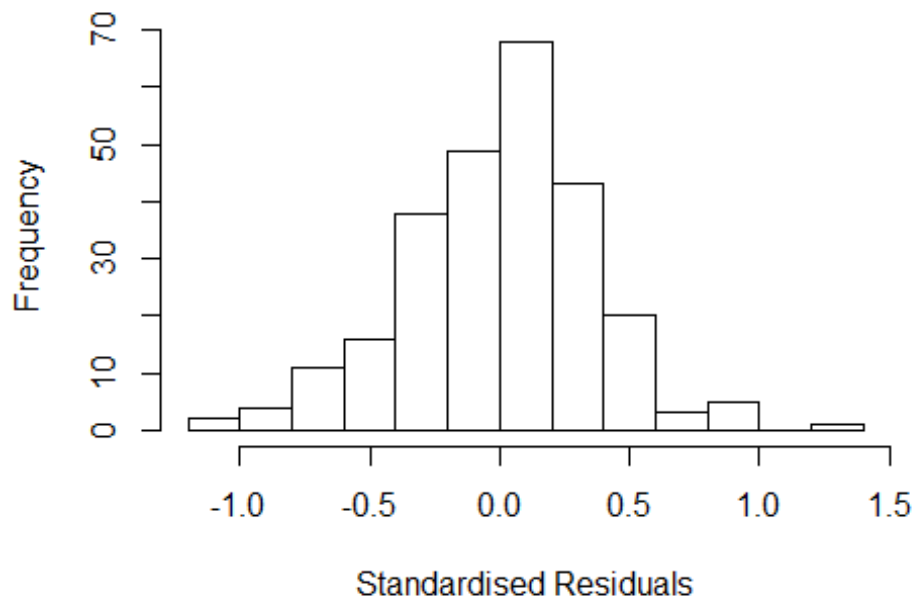
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.983 -0.221  0.014  0.230  1.242
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.20908    0.12846   9.41   <2e-16 ***
## FirstAuthorFemale1 -0.17313    0.06616  -2.62   0.0094 **
## LastAuthorFemale1  0.14482    0.05818   2.49   0.0135 *
## Year1997        -0.21758    0.13848  -1.57   0.1174
## Year1998         0.21915    0.16549   1.32   0.1867
## Year1999         0.04217    0.17645   0.24   0.8113
## Year2000         0.23167    0.18946   1.22   0.2226
## Year2001        -0.05703    0.15597  -0.37   0.7150
## Year2002        -0.08881    0.18678  -0.48   0.6349
## Year2003        -0.15552    0.14789  -1.05   0.2940
## Year2004        -0.00452    0.14989  -0.03   0.9760
## Year2005        -0.05184    0.15420  -0.34   0.7370
```

```

## Year2006          -0.02974      0.15374      -0.19      0.8468
## Year2007          -0.08880      0.15060      -0.59      0.5560
## Year2008          -0.12588      0.14658      -0.86      0.3913
## Year2009          -0.33570      0.21130      -1.59      0.1134
## Year2010           0.14385      0.15572       0.92      0.3565
## Year2011           0.08476      0.14240       0.60      0.5522
## Year2012           0.06907      0.15180       0.46      0.6495
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.353
## Multiple R-squared:  0.14,   Adjusted R-squared:  0.0757
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 237 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.190  0.888  0.955  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      3.85e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.466 1      1.211
## Year              1.466 16      1.012

```


Residuals from first author



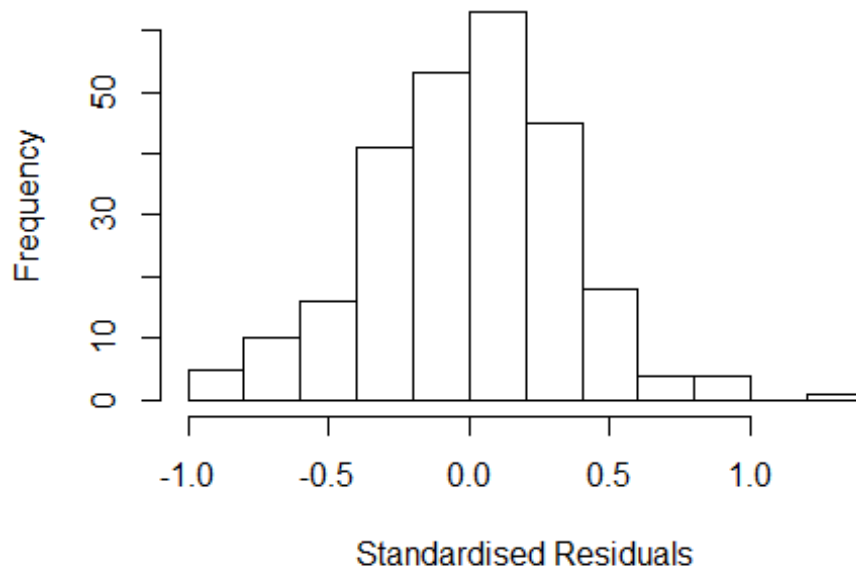
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0023 -0.2172 0.0156 0.2291 1.2907
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2157 0.1218 9.98 <2e-16 ***
## FirstAuthorFemale1 -0.0924 0.0593 -1.56 0.121
## Year1997 -0.2181 0.1319 -1.65 0.099 .
## Year1998 0.2258 0.1665 1.36 0.176
## Year1999 0.0831 0.1728 0.48 0.631
## Year2000 0.2251 0.1844 1.22 0.224
## Year2001 -0.0300 0.1552 -0.19 0.847
## Year2002 -0.0526 0.1824 -0.29 0.773
## Year2003 -0.1471 0.1446 -1.02 0.310
## Year2004 0.0276 0.1434 0.19 0.847
## Year2005 -0.0596 0.1473 -0.40 0.686
## Year2006 -0.0249 0.1517 -0.16 0.870
```

```

## Year2007          -0.0693      0.1425   -0.49    0.627
## Year2008          -0.1186      0.1395   -0.85    0.396
## Year2009          -0.3270      0.2080   -1.57    0.117
## Year2010           0.1321      0.1491    0.89    0.376
## Year2011           0.0741      0.1356    0.55    0.585
## Year2012           0.0817      0.1492    0.55    0.585
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.122, Adjusted R-squared:  0.0602
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 235 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.180  0.885   0.955   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      3.85e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.282 1          1.132
## Year              1.282 16          1.008

```

Residuals from last author



```
## [1] "List of 0 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
## -0.9722 -0.2435 0.0238 0.2138 1.2116
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1863 0.1258 9.43 <2e-16 ***
## LastAuthorFemale1 0.0570 0.0520 1.10 0.274
## Year1997 -0.2027 0.1372 -1.48 0.141
## Year1998 0.2239 0.1644 1.36 0.174
## Year1999 0.0838 0.1745 0.48 0.631
## Year2000 0.2544 0.1871 1.36 0.175
## Year2001 -0.0275 0.1575 -0.17 0.862
## Year2002 -0.0412 0.1874 -0.22 0.826
## Year2003 -0.1420 0.1460 -0.97 0.332
## Year2004 0.0115 0.1495 0.08 0.939
## Year2005 -0.0605 0.1531 -0.40 0.693
## Year2006 -0.0265 0.1487 -0.18 0.859
```

```

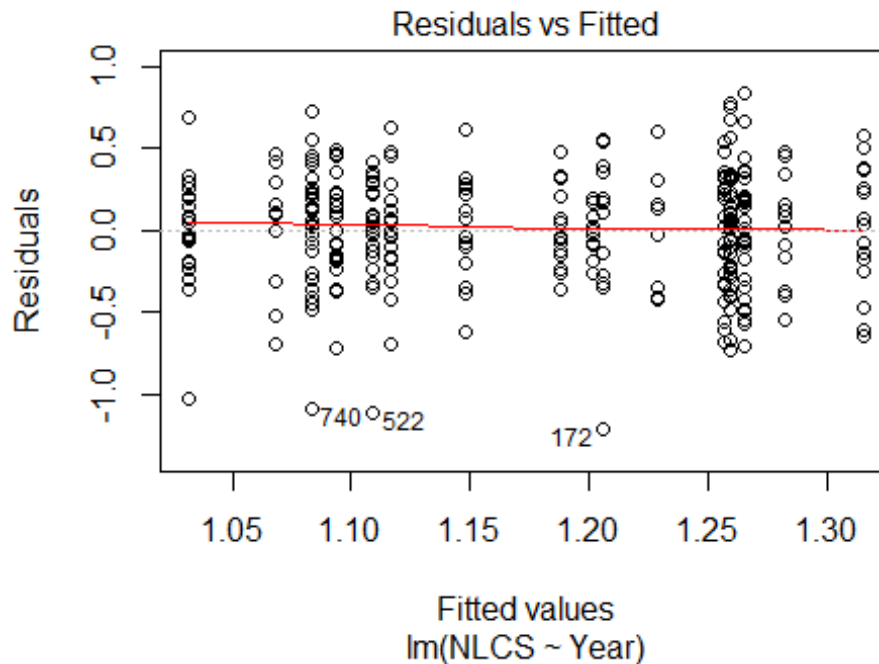
## Year2007          -0.0851      0.1457   -0.58      0.560
## Year2008          -0.1238      0.1435   -0.86      0.389
## Year2009          -0.3679      0.2014   -1.83      0.069 .
## Year2010           0.1058      0.1577     0.67      0.503
## Year2011           0.0427      0.1389     0.31      0.759
## Year2012           0.0717      0.1533     0.47      0.640
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.117, Adjusted R-squared:  0.055
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 240 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.247  0.893  0.957  0.910  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      3.85e-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: 260"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    2    4    4   11    5    3    4    1    4    4    2    3    3    3
## 2012
##    4
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    0    2    2    3    2    1    0    0    3    3    2    3    1    3
## 2012

```

```

##      2
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##      1      0      2      2      3      2      1      0      0      2      3      2      3      1      3
## 2012
##      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: 6"
## [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: 27"
## [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
##      28      26      25      28      33      33      32      35      22      15      28      32      32      24      56
## 2011 2012
##      31      46
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      16      14      8      12      14      15      23      24      12      10      22      22      21      17      33
## 2011 2012
##      21      30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      15      13      7      12      14      14      21      22      10      8      16      21      20      13      32
## 2011 2012
##      17      27
## [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

```



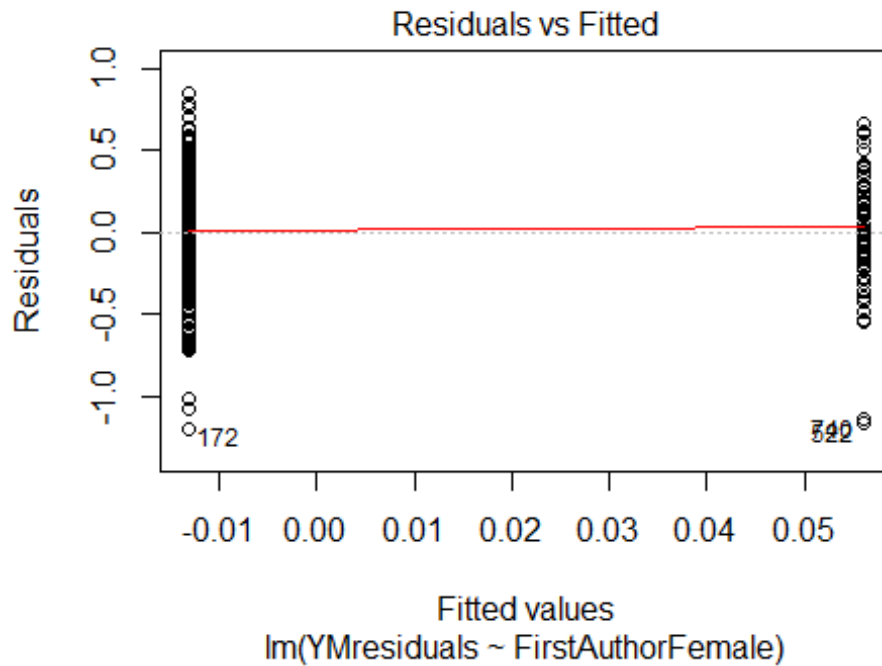
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.79, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 3.62288344132636"
## [1] "Male first author team size 2018 geometric mean: 3.36289489886567"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

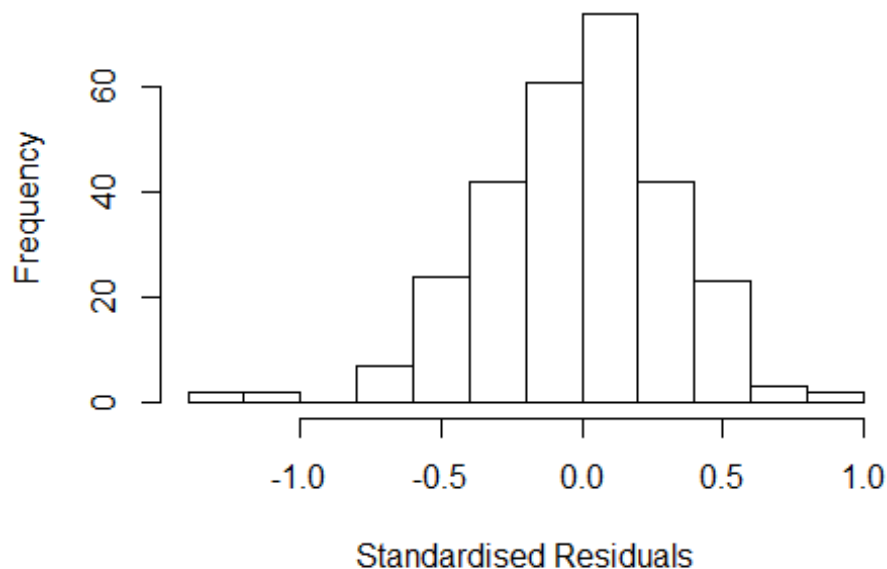
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 55, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.36042145371267"
## [1] "Male last author team size 2018 geometric mean: 3.52891281418207"

## 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 = 38, 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 1.351  1      1.162
## LastAuthorFemale  1.316  1      1.147
## UniqueAuthors    2.633  4      1.129
## Year             3.219 16      1.037
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.2776 -0.2258  0.0103  0.1902  0.8969
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.29565    0.12059   10.74  <2e-16 ***
## FirstAuthorFemale1  0.10039    0.05436    1.85  0.0659 .
## LastAuthorFemale1 -0.05519    0.05973   -0.92  0.3563
## UniqueAuthors2     0.10613    0.06594    1.61  0.1087
## UniqueAuthors3     0.03615    0.06761    0.53  0.5933
## UniqueAuthors4     0.00595    0.08602    0.07  0.9449
## UniqueAuthors5     0.11529    0.09664    1.19  0.2340
## Year1997         -0.08570    0.16429   -0.52  0.6024
## Year1998         -0.15034    0.18782   -0.80  0.4242
## Year1999         -0.05810    0.15717   -0.37  0.7120
```

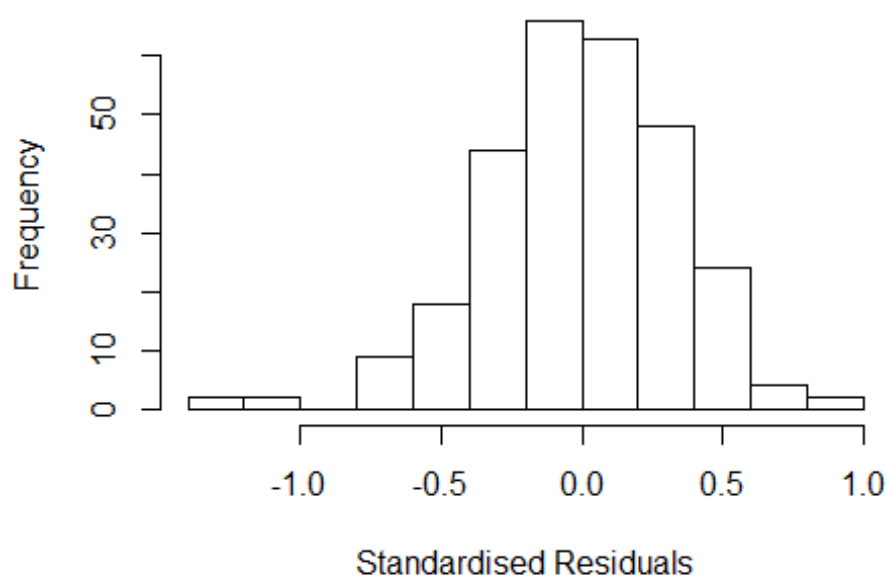


```

## Year2000      -0.04223    0.14070   -0.30    0.7643
## Year2001      -0.21539    0.13849   -1.56    0.1211
## Year2002      -0.07886    0.13859   -0.57    0.5699
## Year2003      -0.11812    0.13226   -0.89    0.3726
## Year2004      -0.18196    0.11771   -1.55    0.1234
## Year2005      -0.25236    0.14856   -1.70    0.0906 .
## Year2006      -0.25623    0.11917   -2.15    0.0325 *
## Year2007      -0.26751    0.13170   -2.03    0.0433 *
## Year2008      -0.33665    0.11983   -2.81    0.0053 **
## Year2009      -0.16214    0.12635   -1.28    0.2005
## Year2010      -0.23368    0.11932   -1.96    0.0513 .
## Year2011      -0.19969    0.12955   -1.54    0.1244
## Year2012      -0.15461    0.13934   -1.11    0.2682
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.339
## Multiple R-squared:  0.0909, Adjusted R-squared:  0.0137
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 261 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.124  0.884  0.957   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      3.55e-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 1.324 1      1.151
## LastAuthorFemale  1.218 1      1.104
## Year              1.367 16      1.010

```

Residuals from first and last author



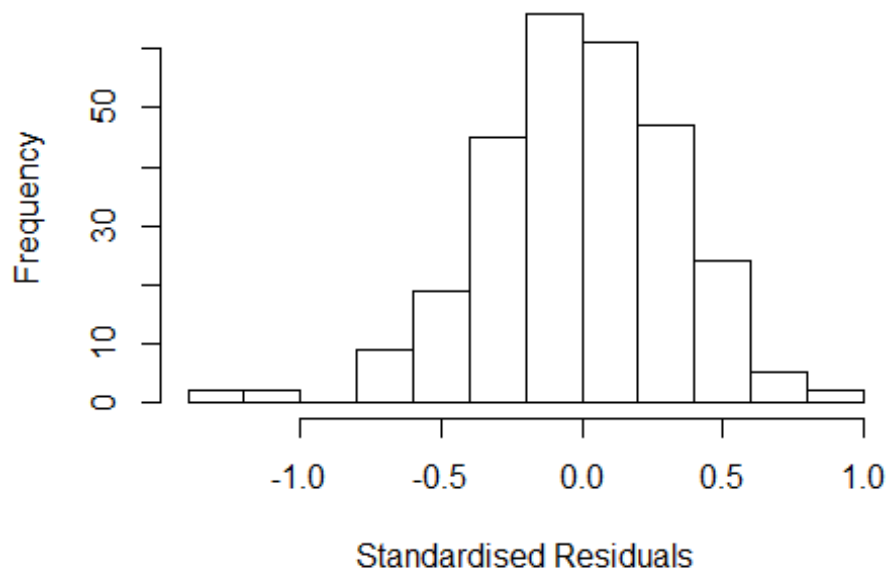
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.29920 -0.22101 -0.00112 0.22071 0.91486
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3485 0.1071 12.59 <2e-16 ***
## FirstAuthorFemale1 0.0968 0.0567 1.71 0.0890 .
## LastAuthorFemale1 -0.0373 0.0578 -0.64 0.5196
## Year1997 -0.1046 0.1601 -0.65 0.5141
## Year1998 -0.1898 0.1813 -1.05 0.2963
## Year1999 -0.0493 0.1688 -0.29 0.7707
## Year2000 -0.0664 0.1361 -0.49 0.6261
## Year2001 -0.2073 0.1392 -1.49 0.1376
## Year2002 -0.0639 0.1406 -0.45 0.6496
## Year2003 -0.1164 0.1291 -0.90 0.3680
## Year2004 -0.1594 0.1167 -1.37 0.1733
## Year2005 -0.2482 0.1534 -1.62 0.1069
```

```

## Year2006          -0.2400      0.1234   -1.94   0.0529 .
## Year2007          -0.2667      0.1312   -2.03   0.0431 *
## Year2008          -0.3335      0.1201   -2.78   0.0059 **
## Year2009          -0.1644      0.1284   -1.28   0.2015
## Year2010          -0.2317      0.1229   -1.89   0.0604 .
## Year2011          -0.1986      0.1281   -1.55   0.1223
## Year2012          -0.1281      0.1388   -0.92   0.3570
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.335
## Multiple R-squared:  0.0755, Adjusted R-squared:  0.0122
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 254 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0983 0.8780 0.9530 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      3.55e-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 1.223 1      1.106
## Year              1.223 16      1.006

```

Residuals from first author



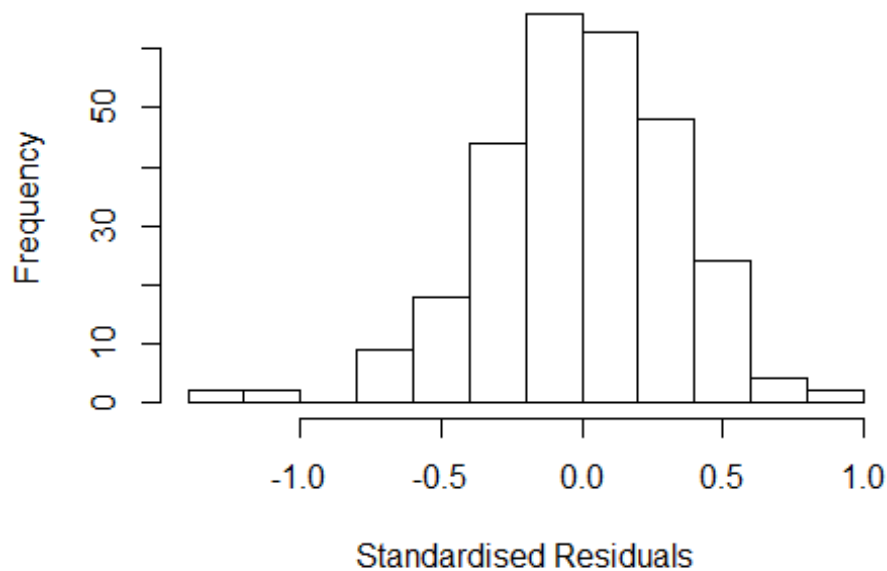
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30028 -0.21990 -0.00445  0.22063  0.88169
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3487    0.1073   12.57  <2e-16 ***
## FirstAuthorFemale1  0.0916    0.0554    1.65  0.0997 .
## Year1997         -0.1043    0.1604   -0.65  0.5160
## Year1998         -0.1969    0.1802   -1.09  0.2755
## Year1999         -0.0484    0.1686   -0.29  0.7744
## Year2000         -0.0717    0.1366   -0.52  0.6003
## Year2001         -0.2098    0.1397   -1.50  0.1344
## Year2002         -0.0672    0.1401   -0.48  0.6316
## Year2003         -0.1184    0.1288   -0.92  0.3588
## Year2004         -0.1661    0.1177   -1.41  0.1594
## Year2005         -0.2475    0.1535   -1.61  0.1081
## Year2006         -0.2470    0.1227   -2.01  0.0451 *
```

```

## Year2007          -0.2743      0.1301   -2.11   0.0360 *
## Year2008          -0.3363      0.1200   -2.80   0.0054 **
## Year2009          -0.1699      0.1285   -1.32   0.1872
## Year2010          -0.2375      0.1235   -1.92   0.0556 .
## Year2011          -0.2052      0.1283   -1.60   0.1109
## Year2012          -0.1324      0.1396   -0.95   0.3438
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.333
## Multiple R-squared:  0.0743, Adjusted R-squared:  0.0146
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 256 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0932 0.8790 0.9530 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      3.55e-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 1.106 1          1.052
## Year            1.106 16          1.003

```

Residuals from last author



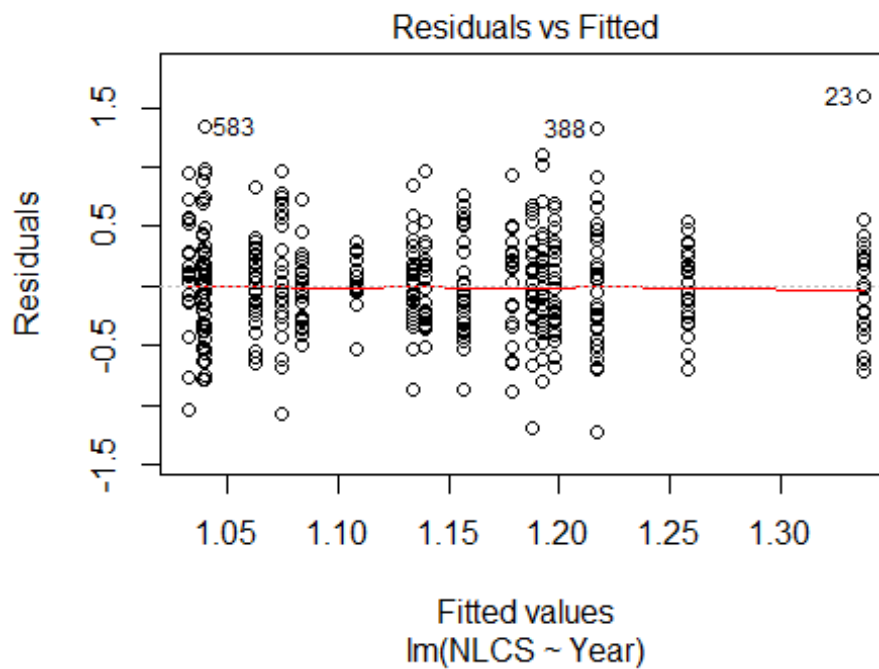
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3053 -0.2197 0.0157 0.2127 0.8668
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3485 0.1072 12.58 <2e-16 ***
## LastAuthorFemale1 -0.0201 0.0558 -0.36 0.7194
## Year1997 -0.0987 0.1598 -0.62 0.5371
## Year1998 -0.1567 0.1961 -0.80 0.4249
## Year1999 -0.0432 0.1642 -0.26 0.7924
## Year2000 -0.0607 0.1370 -0.44 0.6582
## Year2001 -0.1925 0.1395 -1.38 0.1687
## Year2002 -0.0428 0.1407 -0.30 0.7612
## Year2003 -0.1065 0.1280 -0.83 0.4060
## Year2004 -0.1543 0.1154 -1.34 0.1824
## Year2005 -0.2332 0.1527 -1.53 0.1279
## Year2006 -0.2298 0.1233 -1.86 0.0634 .
```

```

## Year2007          -0.2508      0.1292    -1.94    0.0532 .
## Year2008          -0.3141      0.1204    -2.61    0.0096 **
## Year2009          -0.1474      0.1281    -1.15    0.2510
## Year2010          -0.2151      0.1255    -1.71    0.0878 .
## Year2011          -0.1827      0.1273    -1.44    0.1524
## Year2012          -0.0972      0.1364    -0.71    0.4767
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.334
## Multiple R-squared:  0.0646, Adjusted R-squared:  0.00441
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0935 0.8770 0.9530 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      3.55e-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: 282"
## [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
##   55   46   43   51   61   63   42   42   62   50   40   47   31   48   44
## 2011 2012
##   43   57
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   26   29   27   22   30   31   24   27   39   26   29   25   19   27   24
## 2011 2012

```

```
## 19 31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 21 24 25 21 26 29 19 26 34 24 25 22 19 24 22
## 2011 2012
## 17 27
## [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 = 16, p-value = 0.002
```



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

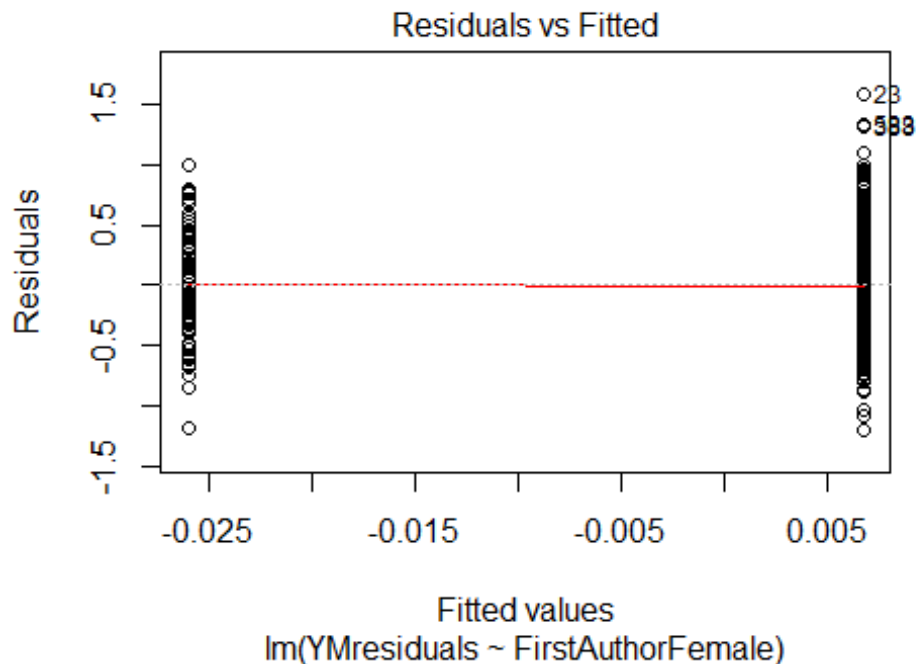
## [1] "Female first author team size 2018 geometric mean: 3.59606099333134"
## [1] "Male first author team size 2018 geometric mean: 2.4594185243188"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



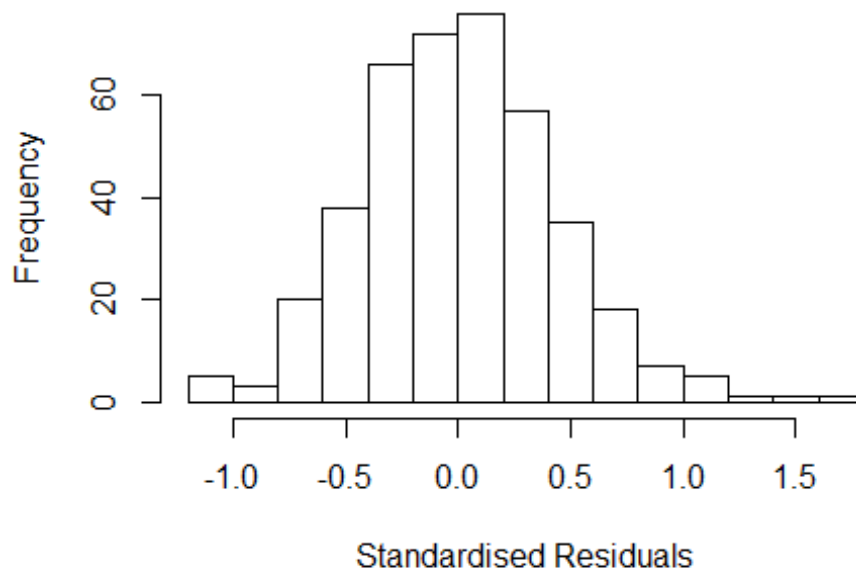
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 320, p-value = 0.01
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.91295063024394"
## [1] "Male last author team size 2018 geometric mean: 2.98041511488199"

## 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 = 77, 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 1.267  1      1.126
## LastAuthorFemale  1.368  1      1.170
## UniqueAuthors    1.960  4      1.088
## Year              2.638 16      1.031
```

Residuals from first and last author and team size



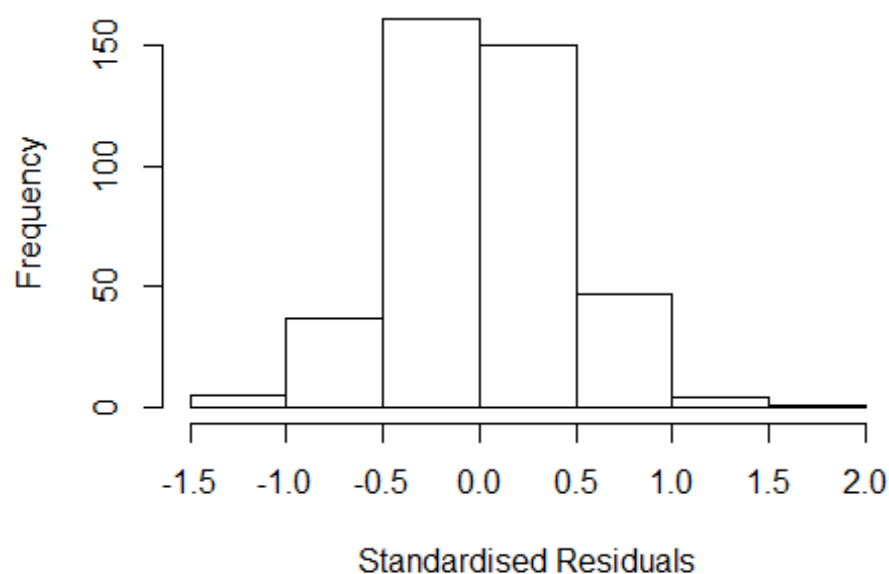
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1723 -0.2781 -0.0045  0.2524  1.6170
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.31203    0.09444   13.89  <2e-16 ***
## FirstAuthorFemale1 -0.04614    0.05256   -0.88  0.3806
## LastAuthorFemale1 -0.00177    0.05335   -0.03  0.9736
## UniqueAuthors2     0.07886    0.06555    1.20  0.2297
## UniqueAuthors3     0.07013    0.06233    1.13  0.2612
## UniqueAuthors4     0.08797    0.07866    1.12  0.2641
## UniqueAuthors5     0.11532    0.07808    1.48  0.1405
## Year1997         -0.19084    0.14236   -1.34  0.1809
## Year1998         -0.22366    0.11459   -1.95  0.0517 .
## Year1999         -0.30388    0.15141   -2.01  0.0455 *
```

```

## Year2000      -0.31063    0.12979   -2.39    0.0172 *
## Year2001      -0.18870    0.14112   -1.34    0.1820
## Year2002      -0.13976    0.14350   -0.97    0.3307
## Year2003      -0.29947    0.14889   -2.01    0.0450 *
## Year2004      -0.36817    0.11629   -3.17    0.0017 **
## Year2005      -0.18457    0.12356   -1.49    0.1361
## Year2006      -0.31079    0.11168   -2.78    0.0057 **
## Year2007      -0.28138    0.11122   -2.53    0.0118 *
## Year2008      -0.18863    0.14521   -1.30    0.1947
## Year2009      -0.08250    0.11258   -0.73    0.4641
## Year2010      -0.26245    0.11536   -2.28    0.0235 *
## Year2011      -0.25326    0.10820   -2.34    0.0198 *
## Year2012      -0.20324    0.13527   -1.50    0.1338
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.0548, Adjusted R-squared:  0.0004
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 373 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0705 0.8670 0.9550 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          2.47e-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 1.204 1 1.097
## LastAuthorFemale 1.297 1 1.139
## Year 1.474 16 1.012

```

Residuals from first and last author



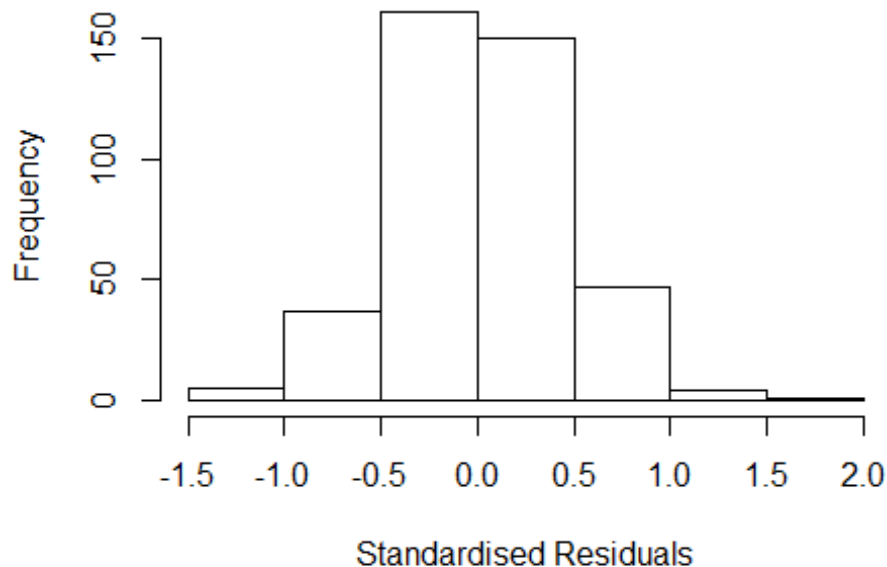
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.252355 -0.275460 -0.000743 0.261904 1.567257
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.36174 0.08257 16.49 <2e-16 ***
## FirstAuthorFemale1 -0.04489 0.05259 -0.85 0.3938
## LastAuthorFemale1 -0.00205 0.05261 -0.04 0.9689
## Year1997 -0.19265 0.13707 -1.41 0.1607
## Year1998 -0.21432 0.10725 -2.00 0.0464 *
## Year1999 -0.30990 0.14596 -2.12 0.0344 *
## Year2000 -0.29890 0.12623 -2.37 0.0184 *
## Year2001 -0.17478 0.13451 -1.30 0.1946
## Year2002 -0.10939 0.13868 -0.79 0.4307
## Year2003 -0.28660 0.14327 -2.00 0.0461 *
## Year2004 -0.35982 0.10994 -3.27 0.0012 **
## Year2005 -0.17487 0.11872 -1.47 0.1416
```

```

## Year2006          -0.29590      0.10382      -2.85      0.0046 **
## Year2007          -0.24863      0.10289      -2.42      0.0161 *
## Year2008          -0.15735      0.13929      -1.13      0.2593
## Year2009          -0.07147      0.10699      -0.67      0.5045
## Year2010          -0.24848      0.11142      -2.23      0.0263 *
## Year2011          -0.23129      0.10107      -2.29      0.0226 *
## Year2012          -0.17470      0.12930      -1.35      0.1774
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.401
## Multiple R-squared:  0.0476, Adjusted R-squared:  0.00321
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 375 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
## 0.0935 0.8640 0.9520 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          2.47e-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 1.166 1          1.080
## Year              1.166 16          1.005

```

Residuals from first author



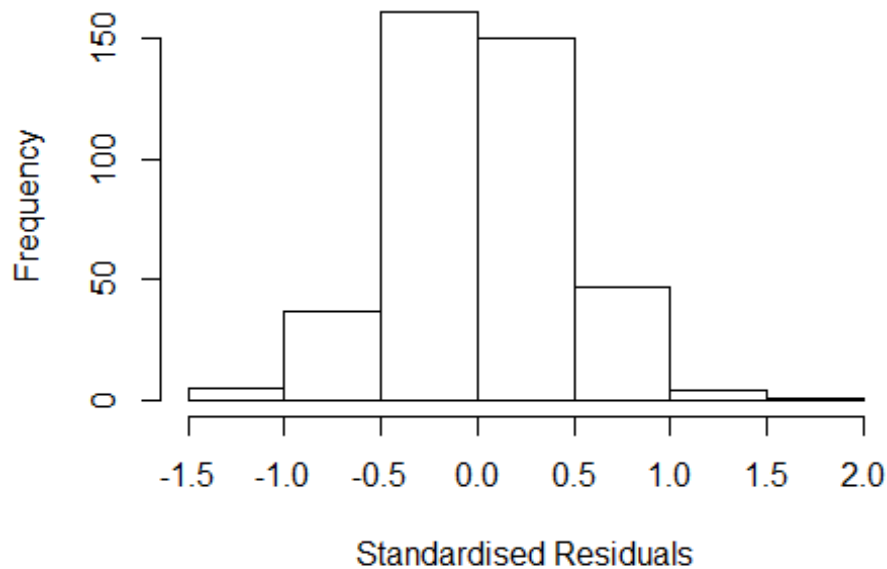
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.251725 -0.276600 -0.000856 0.262232 1.567144
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3619 0.0824 16.53 < 2e-16 ***
## FirstAuthorFemale1 -0.0454 0.0519 -0.88 0.38167
## Year1997 -0.1931 0.1346 -1.43 0.15226
## Year1998 -0.2145 0.1069 -2.01 0.04543 *
## Year1999 -0.3102 0.1457 -2.13 0.03391 *
## Year2000 -0.2996 0.1222 -2.45 0.01467 *
## Year2001 -0.1751 0.1335 -1.31 0.19050
## Year2002 -0.1101 0.1348 -0.82 0.41448
## Year2003 -0.2869 0.1422 -2.02 0.04427 *
## Year2004 -0.3602 0.1082 -3.33 0.00096 ***
## Year2005 -0.1752 0.1179 -1.49 0.13801
## Year2006 -0.2963 0.1027 -2.88 0.00414 **
```

```

## Year2007          -0.2488      0.1023   -2.43  0.01550 *
## Year2008          -0.1575      0.1391   -1.13  0.25813
## Year2009          -0.0717      0.1063   -0.67  0.50024
## Year2010          -0.2490      0.1089   -2.29  0.02272 *
## Year2011          -0.2317      0.0998   -2.32  0.02078 *
## Year2012          -0.1754      0.1250   -1.40  0.16120
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.402
## Multiple R-squared:  0.0476, Adjusted R-squared:  0.00578
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 375 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.094  0.864  0.952  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      2.47e-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 1.256 1          1.121
## Year              1.256 16          1.007

```

Residuals from last author



```
## [1] "List of 0 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
## -1.24408 -0.28076 0.00167 0.26612 1.57690
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3521 0.0822 16.44 <2e-16 ***
## LastAuthorFemale1 -0.0148 0.0526 -0.28 0.7784
## Year1997 -0.1872 0.1377 -1.36 0.1746
## Year1998 -0.2139 0.1079 -1.98 0.0482 *
## Year1999 -0.3060 0.1461 -2.09 0.0369 *
## Year2000 -0.2934 0.1276 -2.30 0.0220 *
## Year2001 -0.1693 0.1324 -1.28 0.2017
## Year2002 -0.1080 0.1412 -0.77 0.4447
## Year2003 -0.2819 0.1448 -1.95 0.0522 .
## Year2004 -0.3573 0.1111 -3.22 0.0014 **
## Year2005 -0.1660 0.1184 -1.40 0.1617
## Year2006 -0.2914 0.1040 -2.80 0.0053 **
```

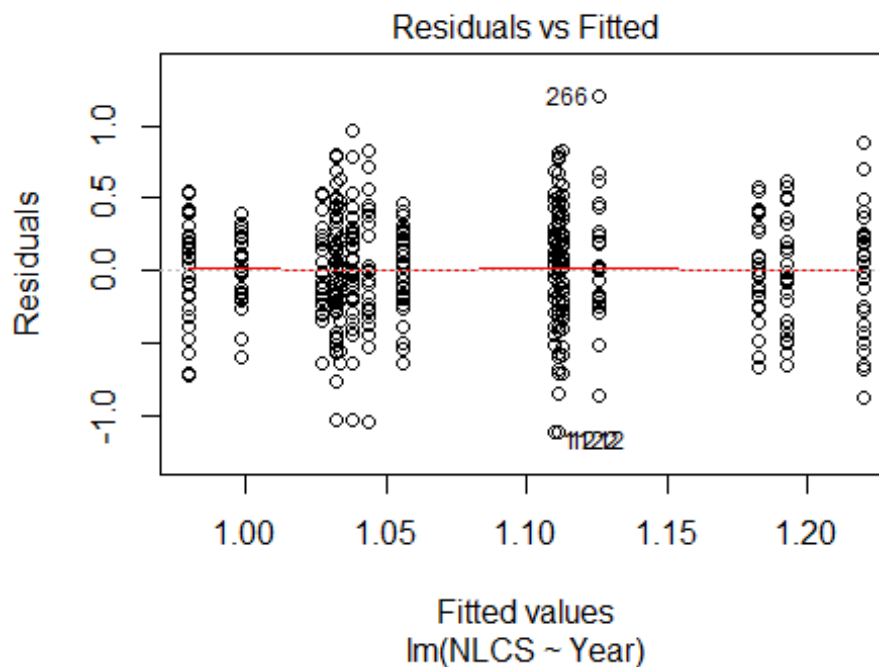


```

## Year2007          -0.2540      0.1039    -2.45    0.0149 *
## Year2008          -0.1611      0.1413    -1.14    0.2547
## Year2009          -0.0735      0.1077    -0.68    0.4956
## Year2010          -0.2454      0.1118    -2.20    0.0287 *
## Year2011          -0.2307      0.1020    -2.26    0.0243 *
## Year2012          -0.1667      0.1288    -1.29    0.1962
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.0459, Adjusted R-squared:  0.00397
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 371 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0939 0.8640 0.9520 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      2.47e-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: 405"
## [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
##   59   51   54   47   66   55   41   63   55   39   49   42   42   35   65
## 2011 2012
##   54   65
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   28   34   29   25   31   25   26   42   40   21   32   26   24   18   30
## 2011 2012

```

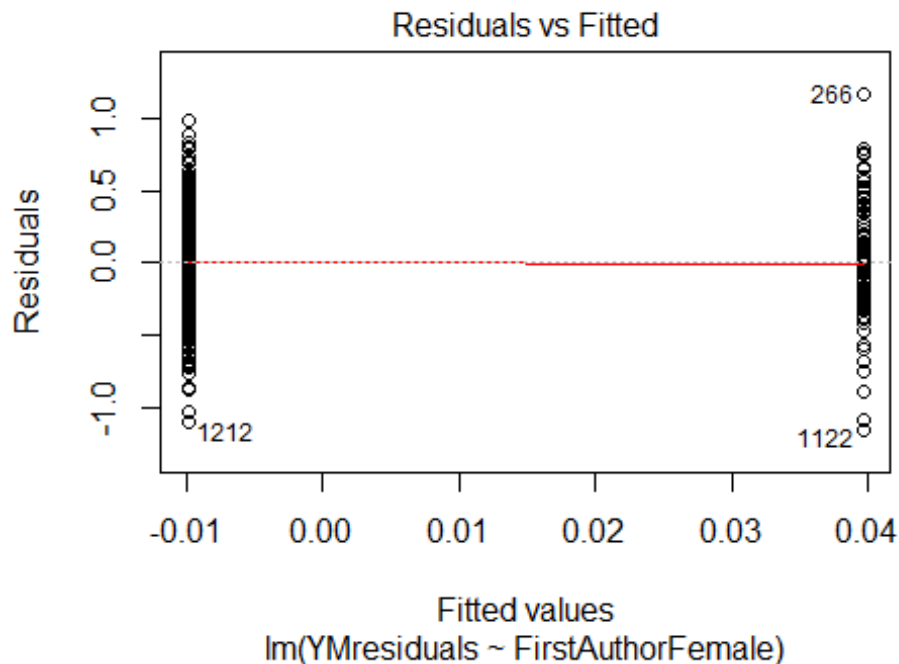
```
## 21 29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 24 28 27 24 29 25 23 40 34 18 25 24 24 16 30
## 2011 2012
## 18 25
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 4.1, df = 1, p-value = 0.04
## [1] "Female first author team size 2018 geometric mean: 4.03190363681863"
## [1] "Male first author team size 2018 geometric mean: 2.84061119116567"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

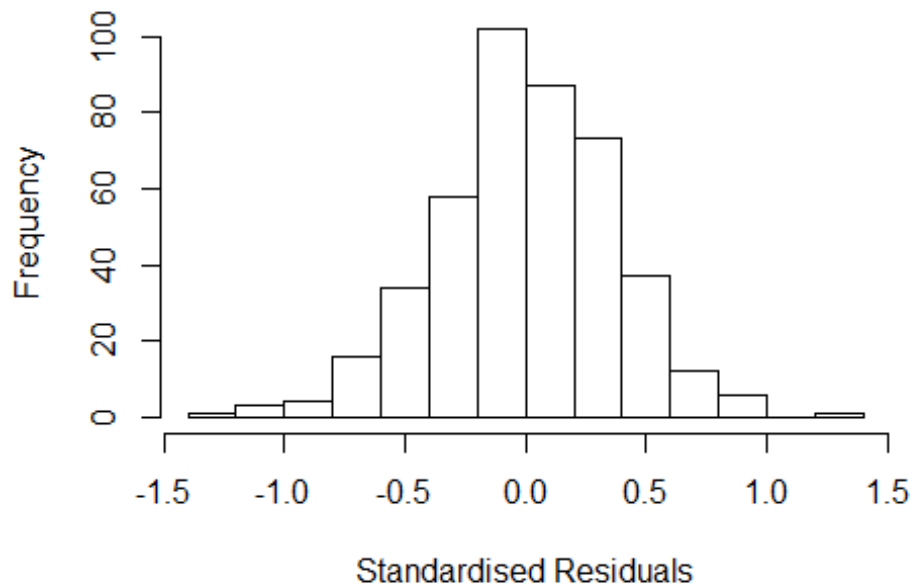
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 260, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.63138857901895"
## [1] "Male last author team size 2018 geometric mean: 3.18270023624117"

## 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 = 0.6
## 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 1.271 1      1.127
## UniqueAuthors    1.900 4      1.084
## Year              2.672 16     1.031
```

Residuals from first and last author and team size



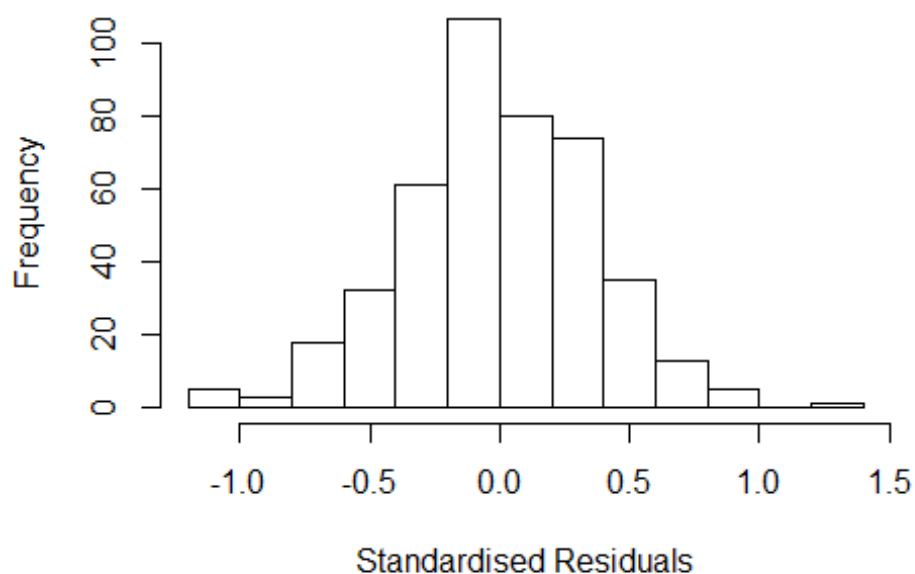
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.30679 -0.21735 -0.00647 0.24143 1.26094
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22285 0.08730 14.01 <2e-16 ***
## FirstAuthorFemale1 0.01928 0.05035 0.38 0.7019
## LastAuthorFemale1 -0.01742 0.05284 -0.33 0.7418
## UniqueAuthors2 0.03183 0.05175 0.62 0.5389
## UniqueAuthors3 0.05791 0.05605 1.03 0.3021
## UniqueAuthors4 -0.02855 0.06285 -0.45 0.6499
## UniqueAuthors5 0.15264 0.06914 2.21 0.0278 *
## Year1997 -0.17820 0.11198 -1.59 0.1123
## Year1998 -0.20940 0.11593 -1.81 0.0716 .
## Year1999 -0.15766 0.12383 -1.27 0.2037
```

```

## Year2000      -0.11911    0.12265   -0.97    0.3320
## Year2001      -0.20079    0.11745   -1.71    0.0881 .
## Year2002      -0.07968    0.11447   -0.70    0.4868
## Year2003      -0.23920    0.09761   -2.45    0.0147 *
## Year2004      -0.27752    0.09970   -2.78    0.0056 **
## Year2005      -0.28163    0.11977   -2.35    0.0192 *
## Year2006      -0.24312    0.09714   -2.50    0.0127 *
## Year2007      -0.22774    0.10045   -2.27    0.0239 *
## Year2008      -0.26709    0.09489   -2.81    0.0051 **
## Year2009      -0.20912    0.11738   -1.78    0.0756 .
## Year2010      -0.08800    0.10061   -0.87    0.3823
## Year2011      -0.06458    0.11864   -0.54    0.5865
## Year2012       0.00753    0.11267    0.07    0.9468
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.356
## Multiple R-squared:  0.0768, Adjusted R-squared:  0.0274
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 398 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.149  0.867  0.955   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          2.30e-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 1.211 1          1.101
## LastAuthorFemale 1.270 1          1.127
## Year              1.483 16          1.012

```

Residuals from first and last author



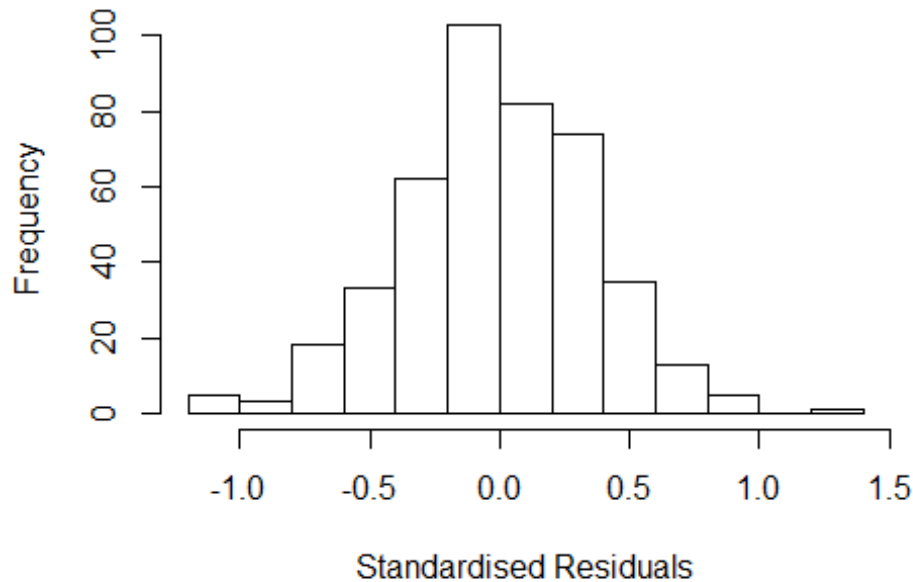
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.183 -0.216 -0.011  0.239  1.236
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2343    0.0795   15.52  <2e-16 ***
## FirstAuthorFemale1  0.0162    0.0519    0.31   0.755
## LastAuthorFemale1 -0.0165    0.0575   -0.29   0.774
## Year1997          -0.1696    0.1121   -1.51   0.131
## Year1998          -0.1955    0.1134   -1.72   0.085 .
## Year1999          -0.1415    0.1219   -1.16   0.246
## Year2000          -0.1017    0.1223   -0.83   0.406
## Year2001          -0.1887    0.1160   -1.63   0.105
## Year2002          -0.0376    0.1113   -0.34   0.736
## Year2003          -0.2074    0.0962   -2.16   0.032 *
## Year2004          -0.2464    0.0974   -2.53   0.012 *
## Year2005          -0.2613    0.1167   -2.24   0.026 *
```

```

## Year2006          -0.2399      0.0981   -2.45    0.015 *
## Year2007          -0.2136      0.0990   -2.16    0.032 *
## Year2008          -0.2301      0.0939   -2.45    0.015 *
## Year2009          -0.2007      0.1212   -1.66    0.099 .
## Year2010          -0.0674      0.1025   -0.66    0.511
## Year2011          -0.0320      0.1238   -0.26    0.796
## Year2012           0.0503      0.1137    0.44    0.659
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.0622, Adjusted R-squared:  0.0215
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 46 weights are ~= 1. The remaining 388 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.207  0.863   0.951   0.901   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      2.30e-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 1.189 1          1.091
## Year              1.189 16          1.005

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.179 -0.219 -0.010  0.240  1.222
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2342    0.0798   15.46  <2e-16 ***
## FirstAuthorFemale1  0.0136    0.0516    0.26  0.7920
## Year1997         -0.1715    0.1120   -1.53  0.1264
## Year1998         -0.1970    0.1132   -1.74  0.0826 .
## Year1999         -0.1417    0.1223   -1.16  0.2475
## Year2000         -0.1068    0.1184   -0.90  0.3676
## Year2001         -0.1911    0.1154   -1.66  0.0986 .
## Year2002         -0.0377    0.1114   -0.34  0.7353
## Year2003         -0.2089    0.0959   -2.18  0.0300 *
## Year2004         -0.2497    0.0956   -2.61  0.0093 **
## Year2005         -0.2611    0.1170   -2.23  0.0262 *
## Year2006         -0.2403    0.0983   -2.44  0.0149 *
```

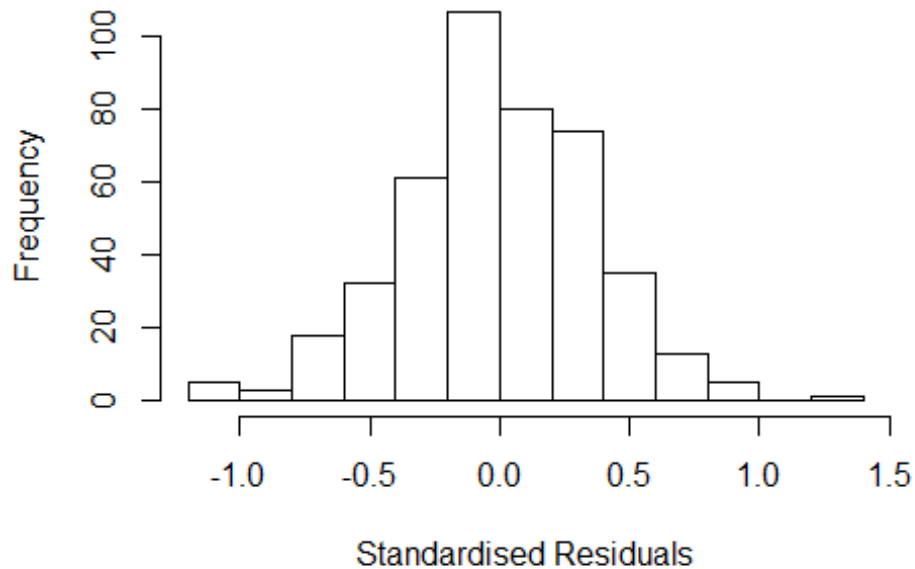


```

## Year2007          -0.2150      0.0988   -2.18   0.0301 *
## Year2008          -0.2321      0.0935   -2.48   0.0135 *
## Year2009          -0.2013      0.1212   -1.66   0.0977 .
## Year2010          -0.0692      0.1026   -0.67   0.5001
## Year2011          -0.0327      0.1242   -0.26   0.7923
## Year2012           0.0491      0.1138    0.43   0.6662
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.356
## Multiple R-squared:  0.0622, Adjusted R-squared:  0.0238
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 45 weights are ~= 1. The remaining 389 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.215  0.864  0.953  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      2.30e-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 1.246 1          1.116
## Year              1.246 16          1.007

```

Residuals from last author



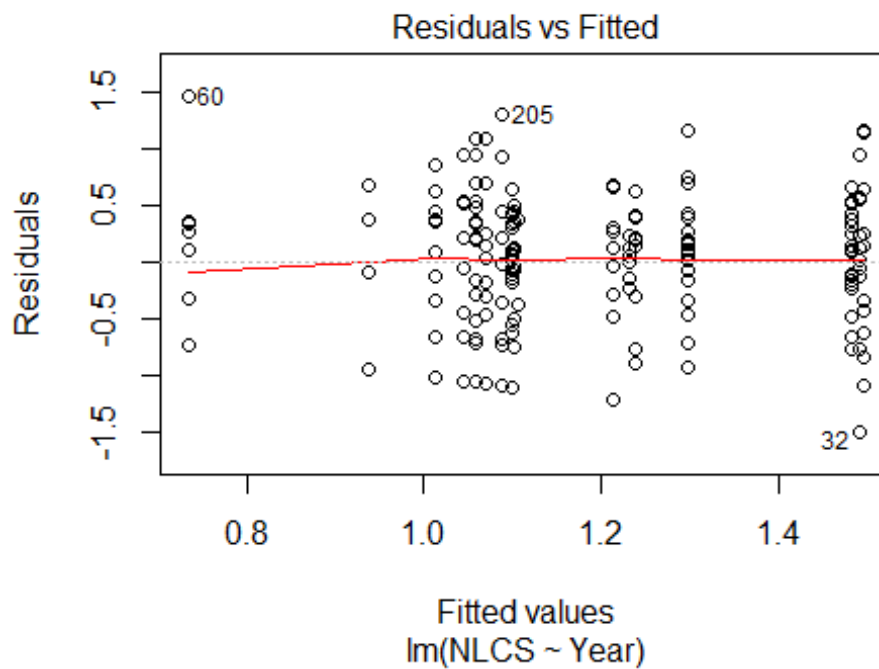
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.16889 -0.21527 -0.00871 0.23732 1.24770
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2360 0.0789 15.67 <2e-16 ***
## LastAuthorFemale1 -0.0132 0.0567 -0.23 0.816
## Year1997 -0.1700 0.1119 -1.52 0.129
## Year1998 -0.1946 0.1138 -1.71 0.088 .
## Year1999 -0.1425 0.1218 -1.17 0.243
## Year2000 -0.1017 0.1219 -0.83 0.405
## Year2001 -0.1886 0.1163 -1.62 0.106
## Year2002 -0.0356 0.1113 -0.32 0.749
## Year2003 -0.2063 0.0962 -2.15 0.033 *
## Year2004 -0.2456 0.0972 -2.53 0.012 *
## Year2005 -0.2619 0.1166 -2.25 0.025 *
## Year2006 -0.2404 0.0980 -2.45 0.015 *
```

```

## Year2007          -0.2112      0.0985    -2.15      0.033 *
## Year2008          -0.2295      0.0941    -2.44      0.015 *
## Year2009          -0.1985      0.1212    -1.64      0.102
## Year2010          -0.0671      0.1034    -0.65      0.517
## Year2011          -0.0307      0.1238    -0.25      0.804
## Year2012           0.0536      0.1139      0.47      0.638
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.062, Adjusted R-squared:  0.0236
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 48 weights are ~= 1. The remaining 386 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.196  0.862  0.952   0.901  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      2.30e-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: 434"
## [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
##   17   25   27   14   14   31   23   11   14   22   13   15   19   32   24
## 2011 2012
##   30   31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2   10    9    4    6   10   10    8    9    9    9   12   13   23   19
## 2011 2012

```

```
## 23 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 2 8 9 4 5 9 8 8 8 9 5 9 12 21 18
## 2011 2012
## 20 14
## [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 = 0.1
```



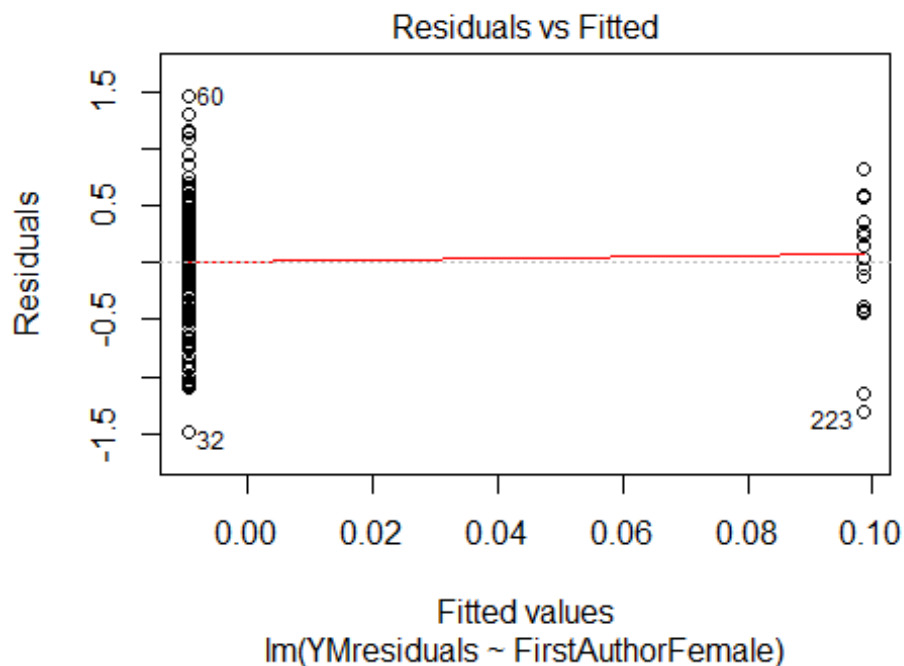
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.14, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 3.76435059950313"
## [1] "Male first author team size 2018 geometric mean: 2.126985641715"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

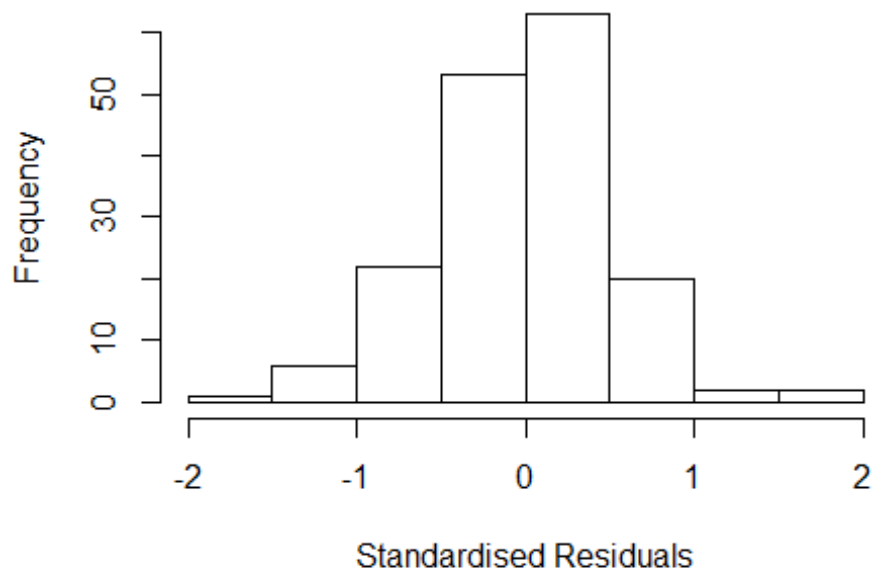
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.008
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 2.57940973180199"

## 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 = 6.5, 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 2.268 1      1.506
## LastAuthorFemale  1.902 1      1.379
## UniqueAuthors    4.027 4      1.190
## Year             10.243 16      1.075
```

Residuals from first and last author and team size



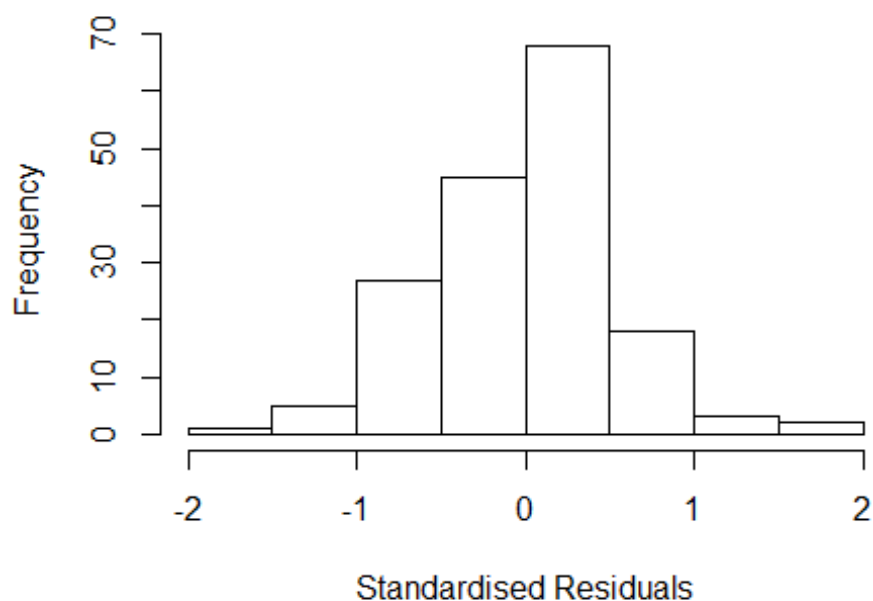
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.522 -0.335  0.025  0.303  1.558
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.12642    0.27550   4.09 7.1e-05 ***
## FirstAuthorFemale1 0.09843    0.13461   0.73  0.4658
## LastAuthorFemale1 0.18610    0.12520   1.49  0.1393
## UniqueAuthors2   -0.04283    0.11406  -0.38  0.7078
## UniqueAuthors3   -0.05801    0.14168  -0.41  0.6828
## UniqueAuthors4    0.24671    0.14075   1.75  0.0817 .
## UniqueAuthors5    0.44680    0.14953   2.99  0.0033 **
## Year1997          0.34455    0.36928   0.93  0.3524
## Year1998         -0.49696    0.40875  -1.22  0.2260
## Year1999         -0.28980    0.37938  -0.76  0.4462
```

```

## Year2000      0.06455      0.28918      0.22      0.8237
## Year2001      0.32209      0.42215      0.76      0.4467
## Year2002     -0.31404      0.33929     -0.93      0.3562
## Year2003     -0.07320      0.35788     -0.20      0.8382
## Year2004     -0.29032      0.40540     -0.72      0.4750
## Year2005      0.11074      0.31240      0.35      0.7235
## Year2006      0.25016      0.34149      0.73      0.4650
## Year2007     -0.01950      0.28993     -0.07      0.9465
## Year2008     -0.24401      0.33801     -0.72      0.4715
## Year2009     -0.00743      0.28055     -0.03      0.9789
## Year2010      0.24457      0.29449      0.83      0.4076
## Year2011      0.16166      0.29161      0.55      0.5802
## Year2012      0.07887      0.33691      0.23      0.8152
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.499
## Multiple R-squared:  0.224, Adjusted R-squared:  0.108
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 146 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.308  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      5.92e-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 2.525 1      1.589
## LastAuthorFemale 1.784 1      1.336
## Year      3.077 16      1.036

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.5013 -0.3278 0.0614 0.3269 1.5842
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10500 0.28714 3.85 0.00018 ***
## FirstAuthorFemale1 0.13036 0.17016 0.77 0.44481
## LastAuthorFemale1 0.13424 0.13591 0.99 0.32489
## Year1997 0.39074 0.37069 1.05 0.29353
## Year1998 -0.50223 0.39331 -1.28 0.20360
## Year1999 -0.23259 0.42741 -0.54 0.58713
## Year2000 0.07500 0.30234 0.25 0.80442
## Year2001 0.38579 0.45538 0.85 0.39824
## Year2002 -0.17138 0.33476 -0.51 0.60944
## Year2003 -0.03868 0.39796 -0.10 0.92271
## Year2004 -0.28308 0.41176 -0.69 0.49283
## Year2005 0.13175 0.33549 0.39 0.69510
```

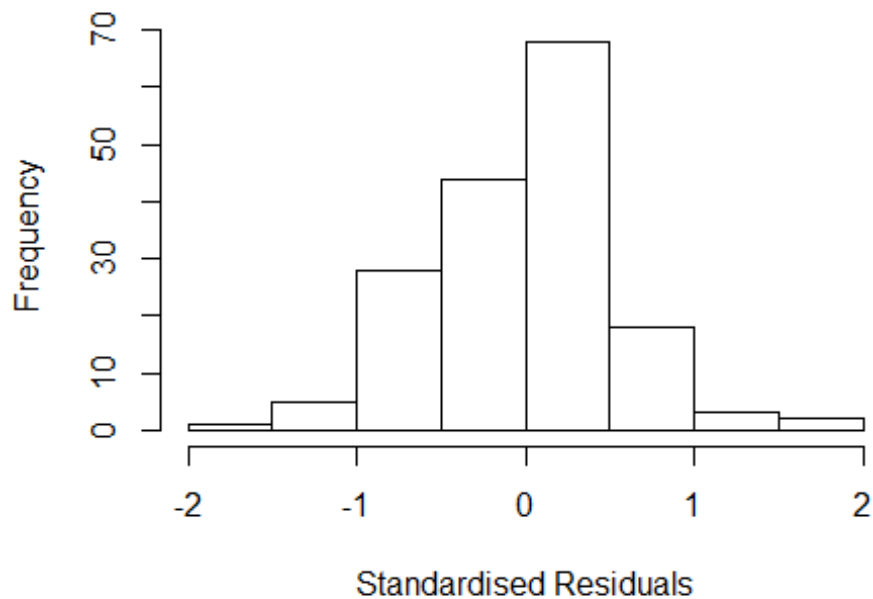


```

## Year2006          0.24230      0.35126      0.69  0.49137
## Year2007          0.11407      0.30642      0.37  0.71023
## Year2008         -0.12397      0.35245     -0.35  0.72552
## Year2009         -0.00994      0.29832     -0.03  0.97347
## Year2010          0.38775      0.30642      1.27  0.20767
## Year2011          0.19800      0.31030      0.64  0.52439
## Year2012          0.09605      0.35098      0.27  0.78473
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.168, Adjusted R-squared:  0.0677
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 158 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.331  0.866  0.952  0.900  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      5.92e-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 1.988 1      1.410
## Year              1.988 16      1.022

```

Residuals from first author



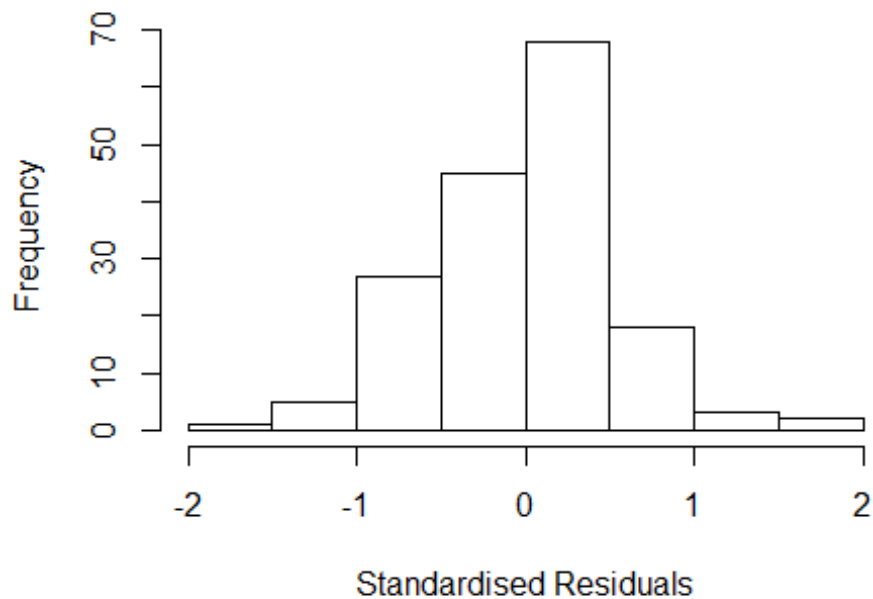
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5383 -0.3458 0.0522 0.3568 1.5735
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1050 0.2876 3.84 0.00018 ***
## FirstAuthorFemale1 0.1832 0.1547 1.18 0.23818
## Year1997 0.4333 0.3893 1.11 0.26751
## Year1998 -0.4915 0.3753 -1.31 0.19227
## Year1999 -0.2177 0.4241 -0.51 0.60851
## Year2000 0.1282 0.2983 0.43 0.66792
## Year2001 0.4038 0.4676 0.86 0.38924
## Year2002 -0.1586 0.3354 -0.47 0.63691
## Year2003 -0.0382 0.4004 -0.10 0.92402
## Year2004 -0.2676 0.4050 -0.66 0.50974
## Year2005 0.1554 0.3409 0.46 0.64926
## Year2006 0.2433 0.3548 0.69 0.49400
```

```

## Year2007          0.1305      0.3057      0.43  0.67014
## Year2008         -0.1124      0.3534     -0.32  0.75099
## Year2009         -0.0154      0.2992     -0.05  0.95910
## Year2010          0.3962      0.3080      1.29  0.20025
## Year2011          0.2038      0.3096      0.66  0.51139
## Year2012          0.0959      0.3535      0.27  0.78655
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.515
## Multiple R-squared:  0.162, Adjusted R-squared:  0.0681
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 154 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.331  0.864  0.953  0.896  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      5.92e-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 1.48 1          1.217
## Year              1.48 16          1.012

```

Residuals from last author



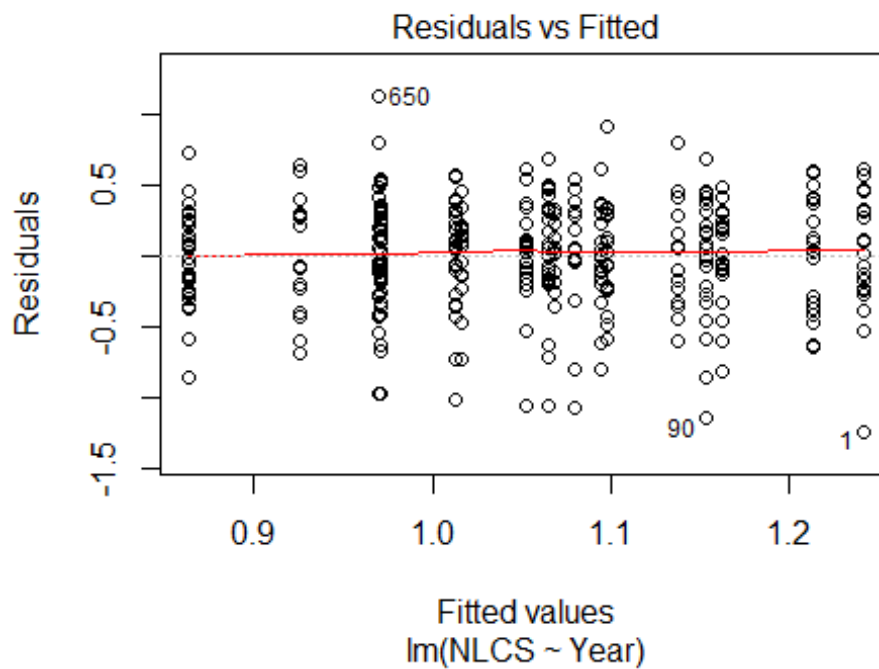
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4719 -0.3479 0.0546 0.3319 1.5619
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1050 0.2848 3.88 0.00016 ***
## LastAuthorFemale1 0.1734 0.1270 1.37 0.17421
## Year1997 0.3669 0.3738 0.98 0.32797
## Year1998 -0.4799 0.3741 -1.28 0.20145
## Year1999 -0.1772 0.4437 -0.40 0.69015
## Year2000 0.0594 0.3001 0.20 0.84332
## Year2001 0.4059 0.4350 0.93 0.35228
## Year2002 -0.1584 0.3326 -0.48 0.63456
## Year2003 -0.0409 0.3932 -0.10 0.91731
## Year2004 -0.2532 0.4019 -0.63 0.52959
## Year2005 0.1208 0.3399 0.36 0.72276
## Year2006 0.2373 0.3519 0.67 0.50108
```

```

## Year2007          0.1091      0.3045      0.36  0.72065
## Year2008         -0.1177      0.3468     -0.34  0.73477
## Year2009          0.0024      0.2959      0.01  0.99354
## Year2010          0.3844      0.3041      1.26  0.20817
## Year2011          0.2012      0.3093      0.65  0.51644
## Year2012          0.1260      0.3390      0.37  0.71068
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.54
## Multiple R-squared:  0.159, Adjusted R-squared:  0.0638
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 155 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.383  0.870  0.957  0.906  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      5.92e-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: 169"
## [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
##   32   30   38   46   29   49   32   27   28   26   30   38   31   34   50
## 2011 2012
##   39   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   19   14   25   27   13   15   22   17   22   15   20   24   17   17   37
## 2011 2012

```

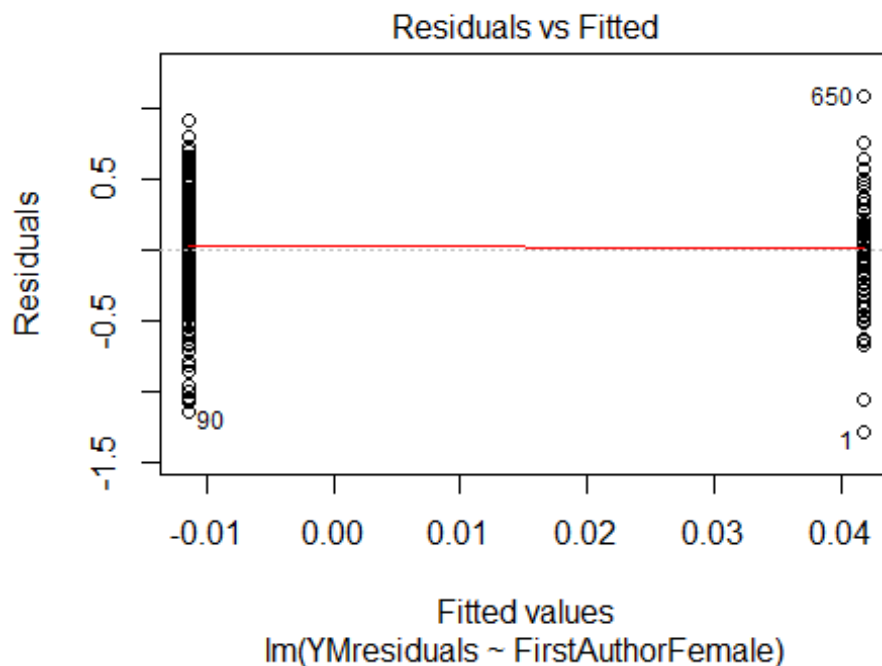
```
## 34 33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 16 14 21 26 11 15 21 16 21 13 19 23 15 17 32
## 2011 2012
## 32 32
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.086, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 1.76273438326762"
## [1] "Male first author team size 2018 geometric mean: 1.93078203668652"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

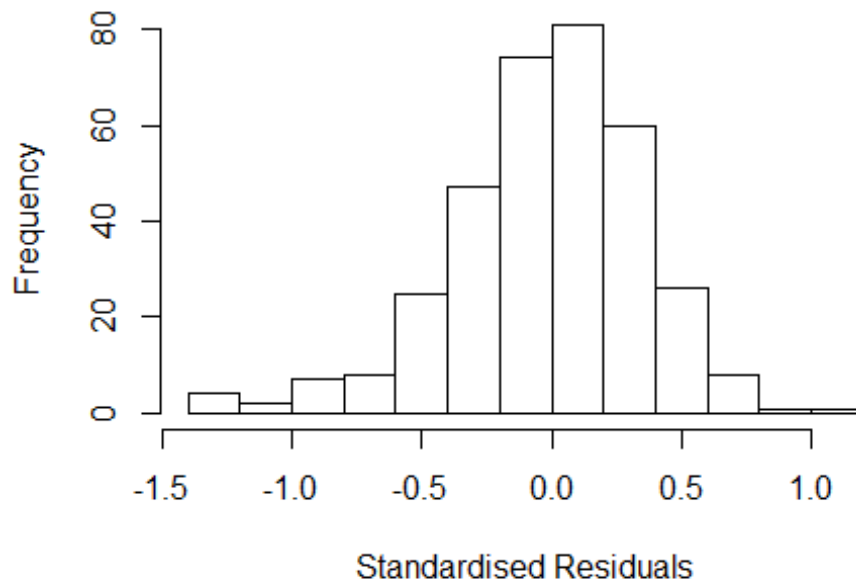
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 28, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male last author team size 2018 geometric mean: 2.20361833090534"

## 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 = 0.2
## 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 1.523  1      1.234
## LastAuthorFemale  1.569  1      1.253
## UniqueAuthors    3.213  4      1.157
## Year              5.294 16      1.053
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.3006 -0.2240  0.0151  0.2267  1.0199
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.18394    0.13314   8.89  <2e-16 ***
## FirstAuthorFemale1  0.06295    0.05380   1.17  0.2428
## LastAuthorFemale1  0.02859    0.05066   0.56  0.5730
## UniqueAuthors2     0.12000    0.06447   1.86  0.0636 .
## UniqueAuthors3     0.05801    0.05983   0.97  0.3330
## UniqueAuthors4     0.17820    0.06920   2.58  0.0105 *
## UniqueAuthors5     0.18451    0.06763   2.73  0.0067 **
## Year1997         -0.10537    0.16669  -0.63  0.5278
## Year1998         -0.00335    0.14659  -0.02  0.9818
## Year1999         -0.13238    0.15261  -0.87  0.3864
```



```

## Year2000      -0.19744    0.13720   -1.44    0.1511
## Year2001      -0.13676    0.16106   -0.85    0.3964
## Year2002      -0.01690    0.15408   -0.11    0.9127
## Year2003      -0.19258    0.14533   -1.33    0.1861
## Year2004      -0.12411    0.14662   -0.85    0.3979
## Year2005      -0.39245    0.18260   -2.15    0.0324 *
## Year2006      -0.17960    0.14169   -1.27    0.2059
## Year2007      -0.20686    0.13711   -1.51    0.1324
## Year2008      -0.28473    0.12752   -2.23    0.0263 *
## Year2009      -0.31739    0.16862   -1.88    0.0607 .
## Year2010      -0.34605    0.13340   -2.59    0.0099 **
## Year2011      -0.25614    0.13583   -1.89    0.0602 .
## Year2012      -0.42247    0.13417   -3.15    0.0018 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.131, Adjusted R-squared:  0.0718
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 316 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.121  0.876  0.954  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      2.91e-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 1.447  1          1.203
## LastAuthorFemale  1.503  1          1.226
## Year              1.712 16          1.017

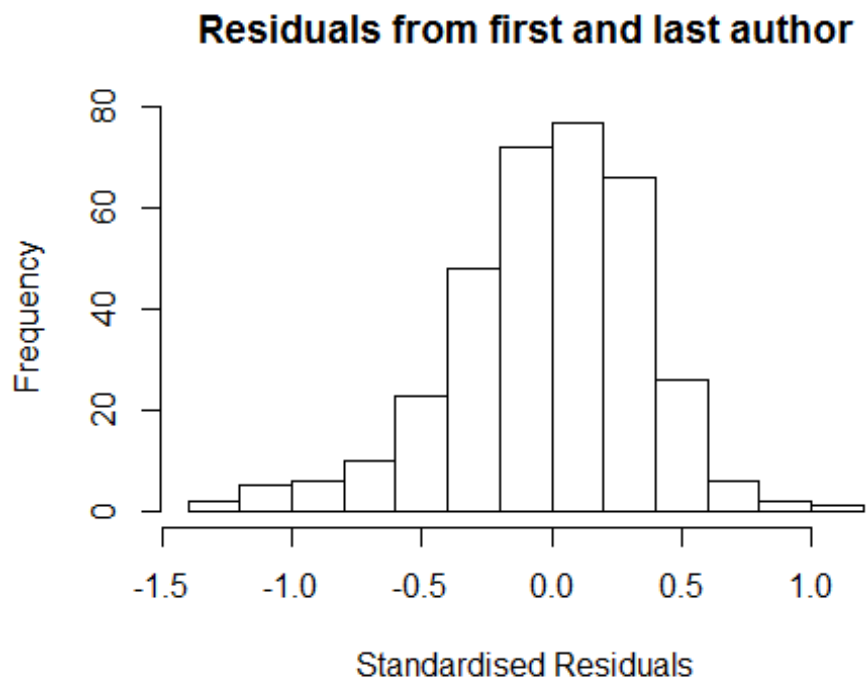
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)

```

```
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37901 -0.21998  0.00843  0.23916  1.09783
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2977     0.1153   11.25 < 2e-16 ***
## FirstAuthorFemale1  0.0518     0.0534    0.97  0.33246
## LastAuthorFemale1  0.0295     0.0520    0.57  0.57037
## Year1997         -0.1771     0.1598   -1.11  0.26864
## Year1998         -0.0580     0.1415   -0.41  0.68229
## Year1999         -0.1951     0.1486   -1.31  0.19015
## Year2000         -0.2503     0.1339   -1.87  0.06244 .
## Year2001         -0.1708     0.1568   -1.09  0.27683
## Year2002         -0.0711     0.1485   -0.48  0.63218
## Year2003         -0.2142     0.1467   -1.46  0.14521
## Year2004         -0.1490     0.1397   -1.07  0.28681
## Year2005         -0.3901     0.1808   -2.16  0.03171 *
## Year2006         -0.2246     0.1358   -1.65  0.09906 .
## Year2007         -0.2367     0.1332   -1.78  0.07646 .
## Year2008         -0.2828     0.1284   -2.20  0.02831 *
## Year2009         -0.3788     0.1628   -2.33  0.02059 *
## Year2010         -0.3489     0.1336   -2.61  0.00943 **
## Year2011         -0.2352     0.1337   -1.76  0.07950 .
## Year2012         -0.4347     0.1294   -3.36  0.00087 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.101, Adjusted R-squared:  0.0517
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 312 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0891  0.8800  0.9540  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      2.91e-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
```

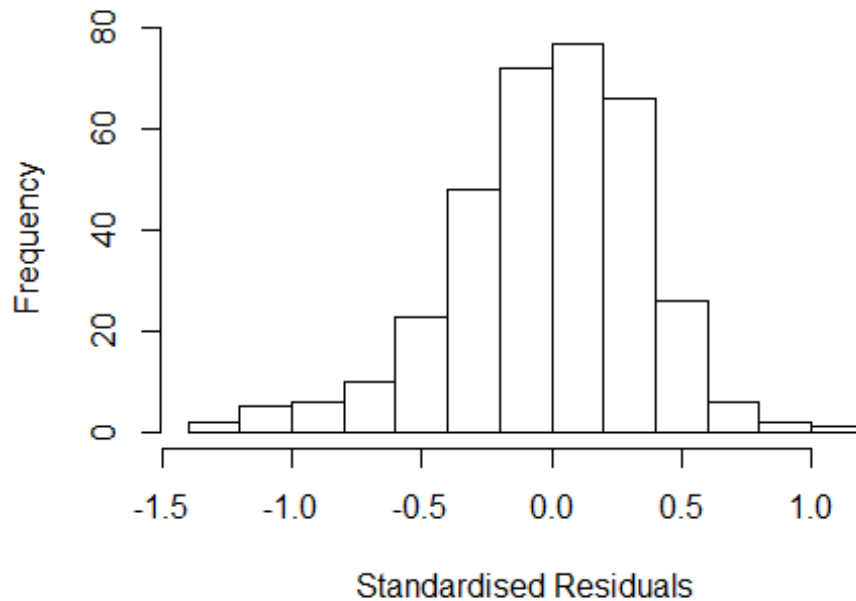
```
##           500           50           2           1           1000           200
## trace.lev      mts compute.rd
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats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

## 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 1.396 1           1.181
## Year             1.396 16           1.010
```

Residuals from last author



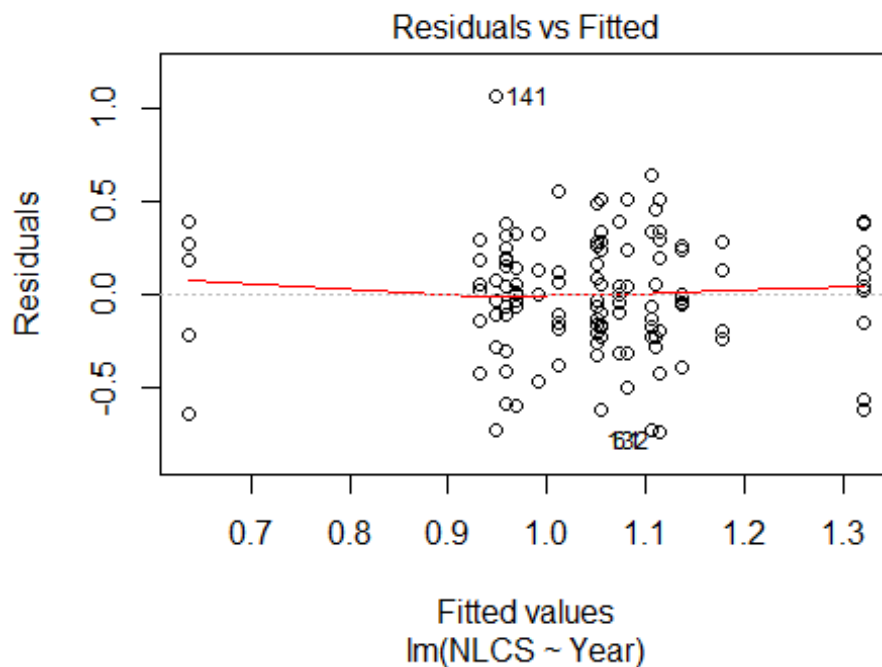
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.34719 -0.21948 0.00535 0.24412 1.12515
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2989 0.1190 10.91 <2e-16 ***
## LastAuthorFemale1 0.0483 0.0499 0.97 0.3342
## Year1997 -0.1790 0.1637 -1.09 0.2752
## Year1998 -0.0483 0.1417 -0.34 0.7333
## Year1999 -0.1951 0.1492 -1.31 0.1919
## Year2000 -0.2500 0.1373 -1.82 0.0697 .
## Year2001 -0.1601 0.1529 -1.05 0.2957
## Year2002 -0.0655 0.1508 -0.43 0.6644
## Year2003 -0.2147 0.1497 -1.43 0.1525
## Year2004 -0.1383 0.1403 -0.99 0.3251
## Year2005 -0.3809 0.1816 -2.10 0.0367 *
## Year2006 -0.2192 0.1390 -1.58 0.1157
```

```

## Year2007          -0.2362      0.1356   -1.74   0.0825 .
## Year2008          -0.2725      0.1314   -2.07   0.0389 *
## Year2009          -0.3743      0.1677   -2.23   0.0263 *
## Year2010          -0.3412      0.1359   -2.51   0.0125 *
## Year2011          -0.2241      0.1349   -1.66   0.0977 .
## Year2012          -0.4285      0.1323   -3.24   0.0013 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.099, Adjusted R-squared:  0.0521
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 315 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.109  0.882  0.954  0.898  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      2.91e-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: 344"
## [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
##   16   18    8    9   14   15    7    7   13   15    8   21   11   10   13
## 2011 2012
##   15    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   10    4    8    4    6    7    6    7   13    4   13    6    5    8
## 2011 2012

```

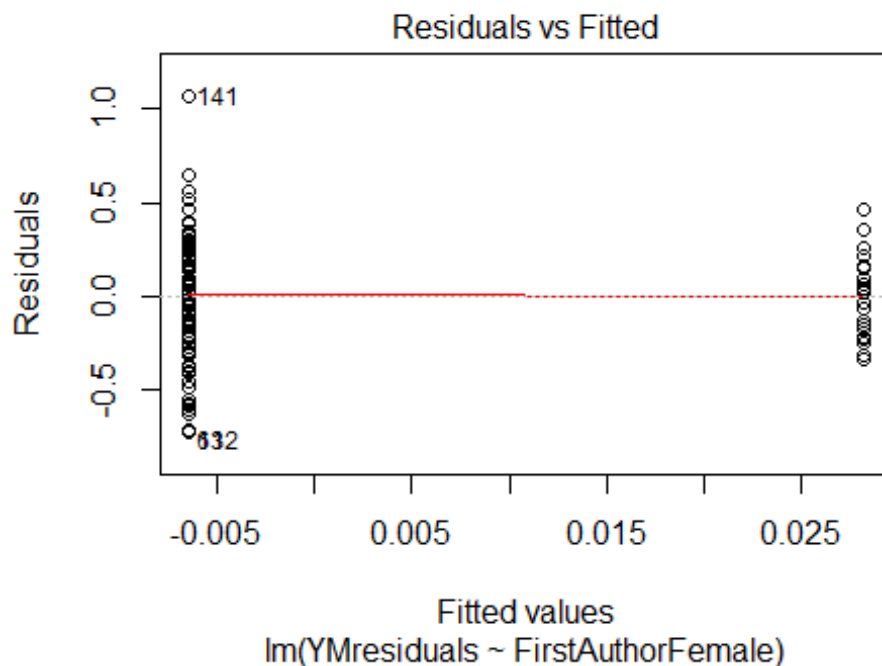
```
## 12 5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 8 4 8 3 6 7 5 5 10 3 11 6 5 7
## 2011 2012
## 12 4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 16, df = 16, p-value = 0.5
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 6.9, df = 1, p-value = 0.009
## [1] "Female first author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male first author team size 2018 geometric mean: 1.57461010625845"
## 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 = 7, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.5874010519682"
## [1] "Male last author team size 2018 geometric mean: 1.5130857494229"

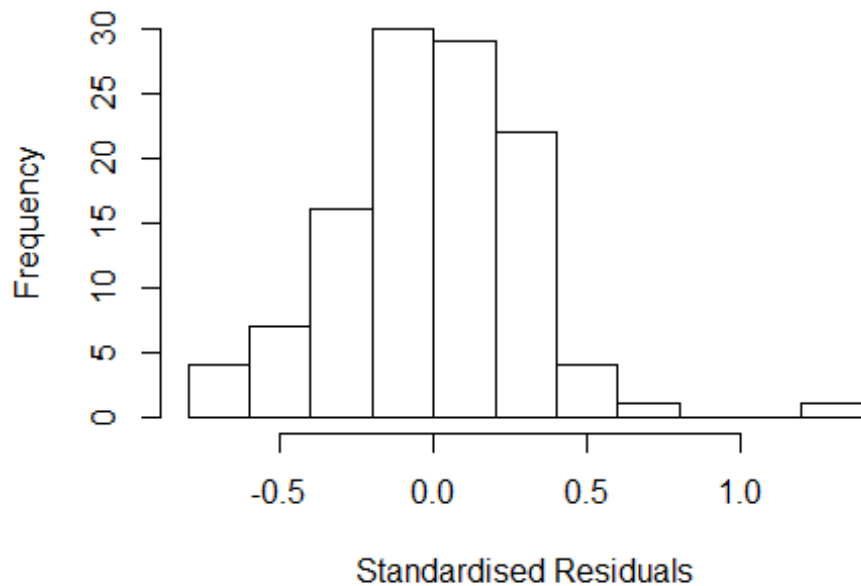
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11, 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	3.412	1	1.847
## LastAuthorFemale	3.652	1	1.911
## UniqueAuthors	11.702	4	1.360
## Year	27.450	16	1.109

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -0.75295 -0.19335 0.00192 0.19101 1.20236
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2830 0.1359 9.44 3.8e-15 ***
## FirstAuthorFemale1 0.0941 0.0880 1.07 0.2880
## LastAuthorFemale1 -0.0768 0.0812 -0.95 0.3468
## UniqueAuthors2 0.1227 0.0999 1.23 0.2227
## UniqueAuthors3 0.0104 0.0862 0.12 0.9045
## UniqueAuthors4 0.0508 0.1099 0.46 0.6452
## UniqueAuthors5 0.2757 0.1019 2.71 0.0081 **
## Year1997 -0.3252 0.1605 -2.03 0.0457 *
## Year1998 -0.2485 0.2046 -1.21 0.2276
## Year1999 -0.1803 0.1951 -0.92 0.3578
```

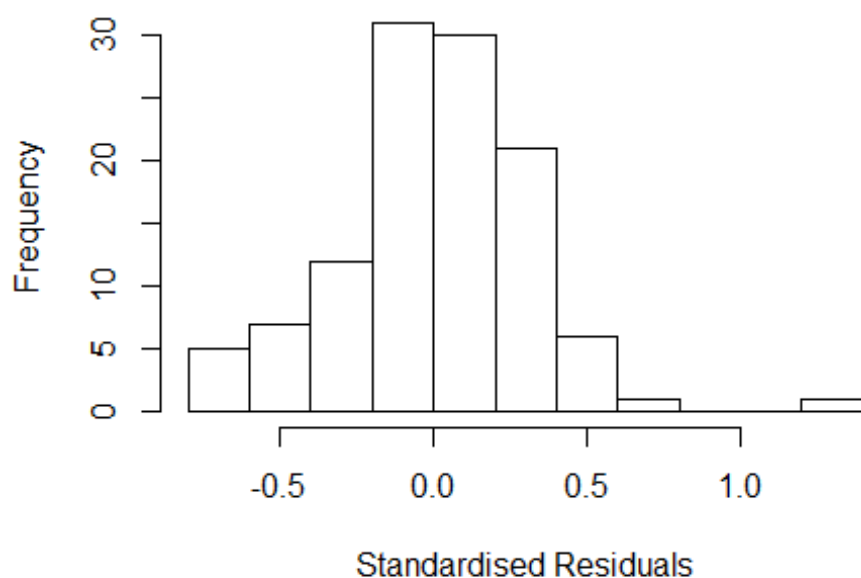


```

## Year2000          -0.2592      0.1287    -2.01    0.0470 *
## Year2001          -0.3683      0.1668    -2.21    0.0297 *
## Year2002          -0.1460      0.2208    -0.66    0.5101
## Year2003          -0.6000      0.2501    -2.40    0.0185 *
## Year2004          -0.2089      0.1472    -1.42    0.1593
## Year2005          -0.3437      0.1325    -2.59    0.0110 *
## Year2006          -0.3541      0.1565    -2.26    0.0260 *
## Year2007          -0.3868      0.1384    -2.79    0.0063 **
## Year2008          -0.4184      0.1491    -2.81    0.0061 **
## Year2009          -0.6664      0.2115    -3.15    0.0022 **
## Year2010          -0.3247      0.1642    -1.98    0.0510 .
## Year2011          -0.4041      0.1366    -2.96    0.0039 **
## Year2012          -0.1745      0.1873    -0.93    0.3540
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.305
## Multiple R-squared:  0.252, Adjusted R-squared:  0.0709
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 102 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0842 0.8910 0.9540 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          8.77e-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 3.145 1 1.773
## LastAuthorFemale 3.774 1 1.943
## Year 3.335 16 1.038

```

Residuals from first and last author



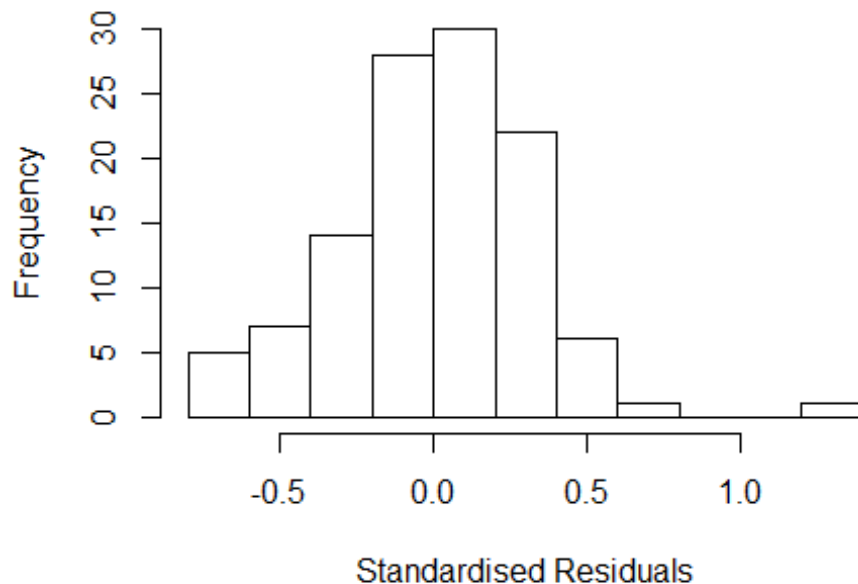
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.79155 -0.17600  0.00805  0.20496  1.26041
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3537     0.1089   12.43  <2e-16 ***
## FirstAuthorFemale1  0.0676     0.0877    0.77   0.4425
## LastAuthorFemale1 -0.0467     0.0878   -0.53   0.5963
## Year1997          -0.3906     0.1322   -2.95   0.0040 **
## Year1998          -0.2806     0.2111   -1.33   0.1869
## Year1999          -0.2343     0.1997   -1.17   0.2437
## Year2000          -0.2072     0.1345   -1.54   0.1268
## Year2001          -0.2915     0.1392   -2.09   0.0389 *
## Year2002          -0.1782     0.2383   -0.75   0.4566
## Year2003          -0.6061     0.2152   -2.82   0.0059 **
## Year2004          -0.2562     0.1432   -1.79   0.0768 .
## Year2005          -0.3320     0.1363   -2.44   0.0167 *
```

```

## Year2006          -0.2602      0.1447    -1.80    0.0754 .
## Year2007          -0.3622      0.1280    -2.83    0.0057 **
## Year2008          -0.4178      0.1447    -2.89    0.0048 **
## Year2009          -0.6875      0.2282    -3.01    0.0033 **
## Year2010          -0.3485      0.1662    -2.10    0.0387 *
## Year2011          -0.2949      0.1436    -2.05    0.0428 *
## Year2012          -0.1595      0.2140    -0.75    0.4579
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.301
## Multiple R-squared:  0.208, Adjusted R-squared:  0.0582
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 99 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0412 0.8830 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      8.77e-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 1.774 1      1.332
## Year              1.774 16      1.018

```

Residuals from first author



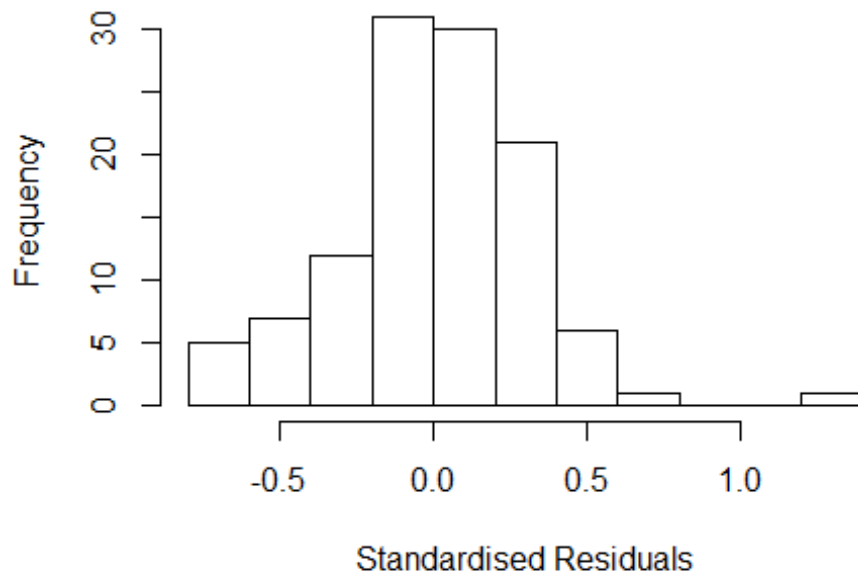
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7895 -0.1713 0.0107 0.2036 1.2889
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3513 0.1089 12.41 <2e-16 ***
## FirstAuthorFemale1 0.0423 0.0685 0.62 0.5388
## Year1997 -0.3888 0.1328 -2.93 0.0043 **
## Year1998 -0.2726 0.2068 -1.32 0.1906
## Year1999 -0.2380 0.1963 -1.21 0.2283
## Year2000 -0.2048 0.1345 -1.52 0.1312
## Year2001 -0.2934 0.1398 -2.10 0.0385 *
## Year2002 -0.1777 0.2341 -0.76 0.4495
## Year2003 -0.6321 0.2196 -2.88 0.0049 **
## Year2004 -0.2726 0.1393 -1.96 0.0532 .
## Year2005 -0.3358 0.1385 -2.42 0.0172 *
## Year2006 -0.2740 0.1488 -1.84 0.0687 .
```

```

## Year2007          -0.3765      0.1278    -2.95    0.0040 **
## Year2008          -0.4193      0.1456    -2.88    0.0049 **
## Year2009          -0.6753      0.2264    -2.98    0.0036 **
## Year2010          -0.3612      0.1666    -2.17    0.0326 *
## Year2011          -0.3037      0.1444    -2.10    0.0380 *
## Year2012          -0.1605      0.2126    -0.76    0.4520
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.297
## Multiple R-squared:  0.207, Adjusted R-squared:  0.0665
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 97 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0197 0.8780 0.9460 0.8950 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      8.77e-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 2.262 1          1.504
## Year            2.262 16          1.026

```

Residuals from last author



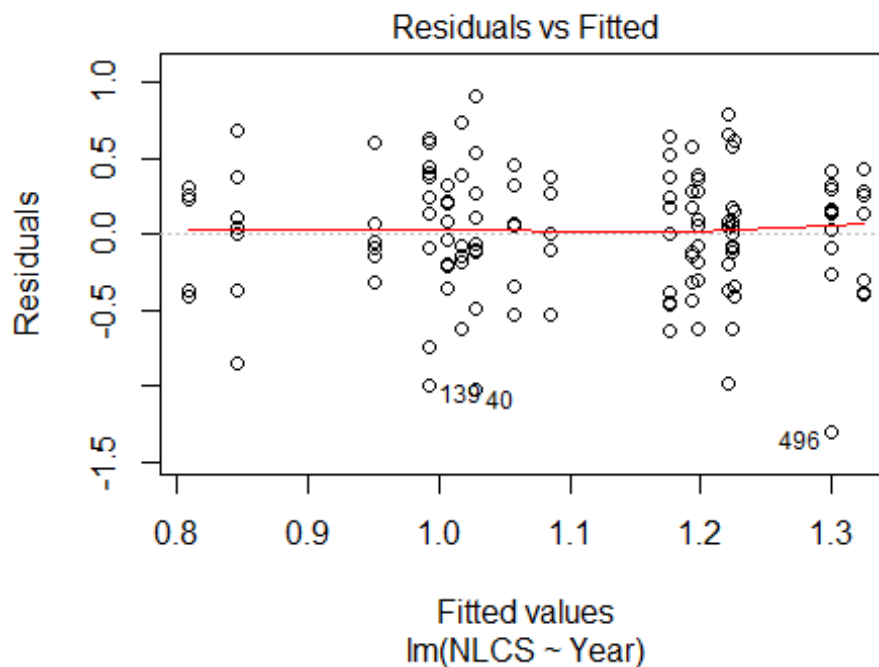
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.79106 -0.17663  0.00728  0.19179  1.28062
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3574     0.1124   12.07  <2e-16 ***
## LastAuthorFemale1 -0.0154     0.0699   -0.22  0.8261
## Year1997         -0.3946     0.1367   -2.89  0.0048 **
## Year1998         -0.2678     0.2015   -1.33  0.1871
## Year1999         -0.2419     0.1992   -1.21  0.2274
## Year2000         -0.2109     0.1375   -1.53  0.1282
## Year2001         -0.2885     0.1418   -2.03  0.0446 *
## Year2002         -0.1823     0.2377   -0.77  0.4450
## Year2003         -0.6300     0.2172   -2.90  0.0046 **
## Year2004         -0.2522     0.1483   -1.70  0.0922 .
## Year2005         -0.3207     0.1405   -2.28  0.0247 *
## Year2006         -0.2748     0.1489   -1.85  0.0681 .
```

```

## Year2007          -0.3667      0.1303    -2.81    0.0059 **
## Year2008          -0.4141      0.1467    -2.82    0.0058 **
## Year2009          -0.6684      0.2268    -2.95    0.0040 **
## Year2010          -0.3278      0.1619    -2.03    0.0456 *
## Year2011          -0.2916      0.1458    -2.00    0.0483 *
## Year2012          -0.1413      0.2002    -0.71    0.4820
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.297
## Multiple R-squared:  0.204, Adjusted R-squared:  0.0629
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 102 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0234 0.8820 0.9530 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      8.77e-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: 114"
## [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
##   14   15   13   15   27   16   12   13   15   19   13   14   15   11   25
## 2011 2012
##   18   24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10    9    4    7   11    7    7    7    6    7    5    7    8    5    9
## 2011 2012

```

```
##      9      11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      8      9      4      7      11      7      7      6      6      6      4      6      8      4      9
## 2011 2012
##      9      10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```

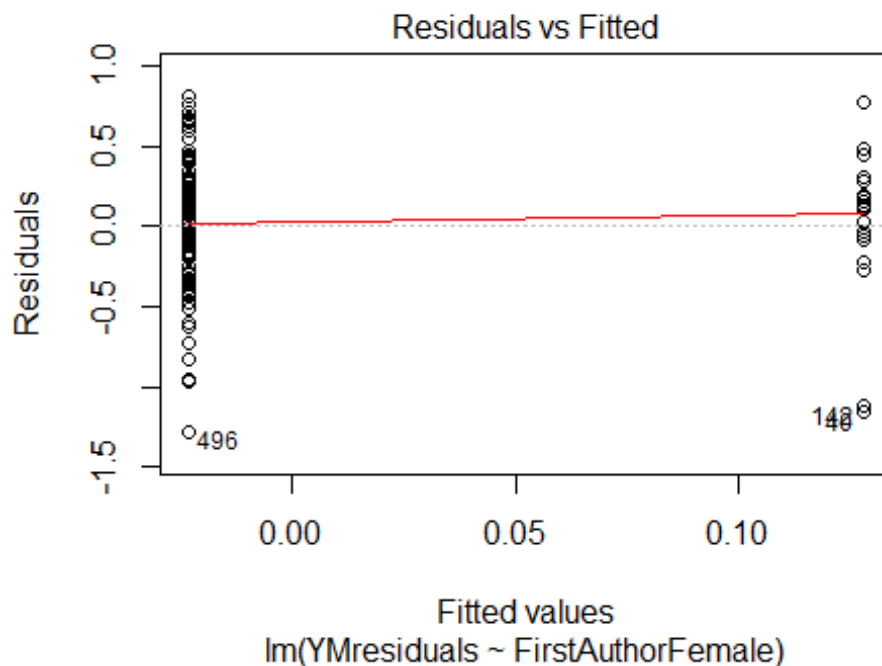


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.83, df = 1, p-value = 0.4
## [1] "Female first author team size 2018 geometric mean: 2.82842712474619"
## [1] "Male first author team size 2018 geometric mean: 1"
## Warning in 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] "Female last author team size 2018 geometric mean: 4"
## [1] "Male last author team size 2018 geometric mean: 2"

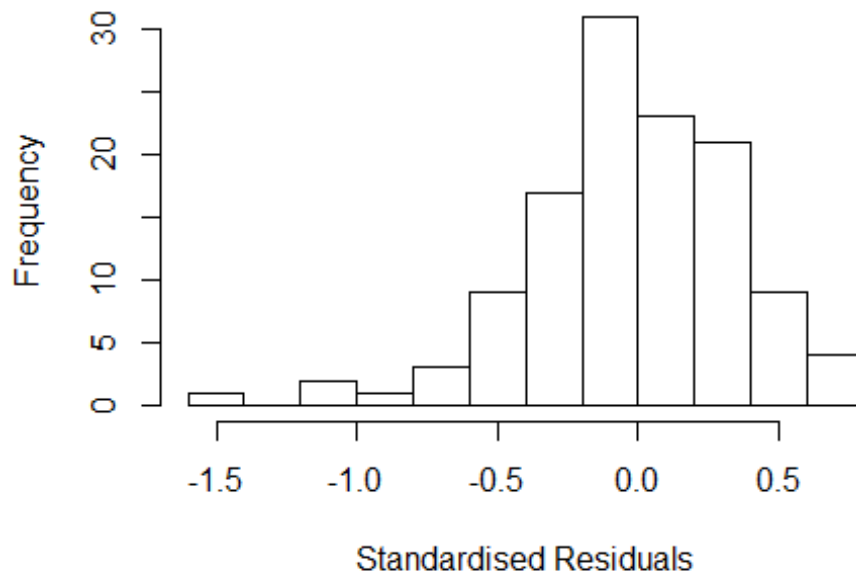
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3.5, 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	3.608	1	1.900
LastAuthorFemale	3.251	1	1.803
UniqueAuthors	23.463	4	1.484
Year	55.612	16	1.134

Residuals from first and last author and team size



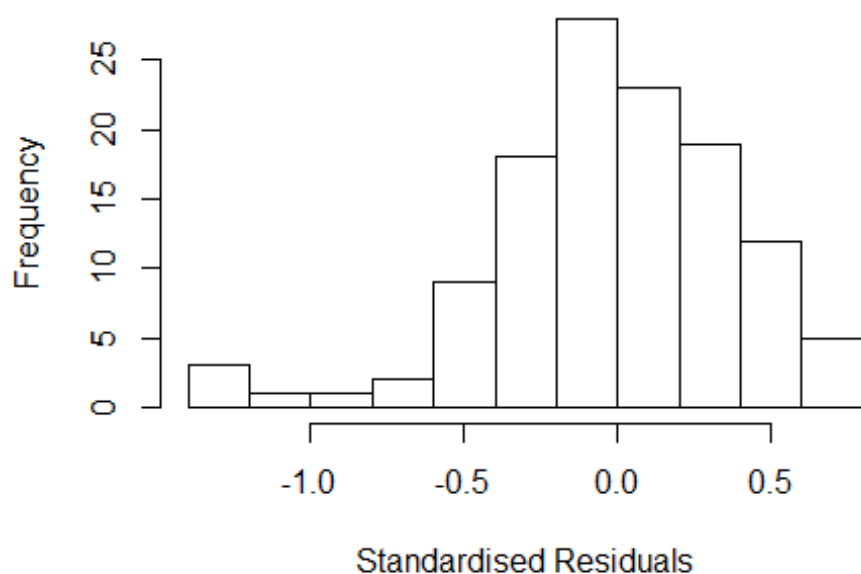
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.42088 -0.23501 -0.00944 0.24501 0.77249
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98356 0.13513 7.28 8.5e-11 ***
## FirstAuthorFemale1 0.23073 0.14182 1.63 0.1070
## LastAuthorFemale1 -0.08885 0.16800 -0.53 0.5981
## UniqueAuthors2 0.31922 0.10071 3.17 0.0020 **
## UniqueAuthors3 0.21678 0.13359 1.62 0.1079
## UniqueAuthors4 0.44249 0.18465 2.40 0.0185 *
## UniqueAuthors5 0.28394 0.17326 1.64 0.1045
## Year1997 0.01150 0.22284 0.05 0.9589
## Year1998 0.07880 0.35696 0.22 0.8258
## Year1999 -0.10809 0.17191 -0.63 0.5310
```

```

## Year2000      -0.00179      0.25775      -0.01      0.9945
## Year2001      0.00966      0.20603      0.05      0.9627
## Year2002      0.07108      0.18964      0.37      0.7086
## Year2003      0.21175      0.29243      0.72      0.4707
## Year2004      -0.07235      0.19560      -0.37      0.7123
## Year2005      -0.05789      0.18168      -0.32      0.7507
## Year2006      -0.30330      0.19878      -1.53      0.1303
## Year2007      -0.20205      0.19773      -1.02      0.3094
## Year2008      -0.25878      0.15251      -1.70      0.0929 .
## Year2009      -0.44678      0.15223      -2.93      0.0042 **
## Year2010      -0.01623      0.14675      -0.11      0.9121
## Year2011      -0.04132      0.15764      -0.26      0.7938
## Year2012      0.05608      0.21479      0.26      0.7946
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.354
## Multiple R-squared:  0.287, Adjusted R-squared:  0.127
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 107 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0705 0.8680 0.9510 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          8.26e-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 2.404 1 1.550
## LastAuthorFemale 1.715 1 1.310
## Year 3.279 16 1.038

```

Residuals from first and last author



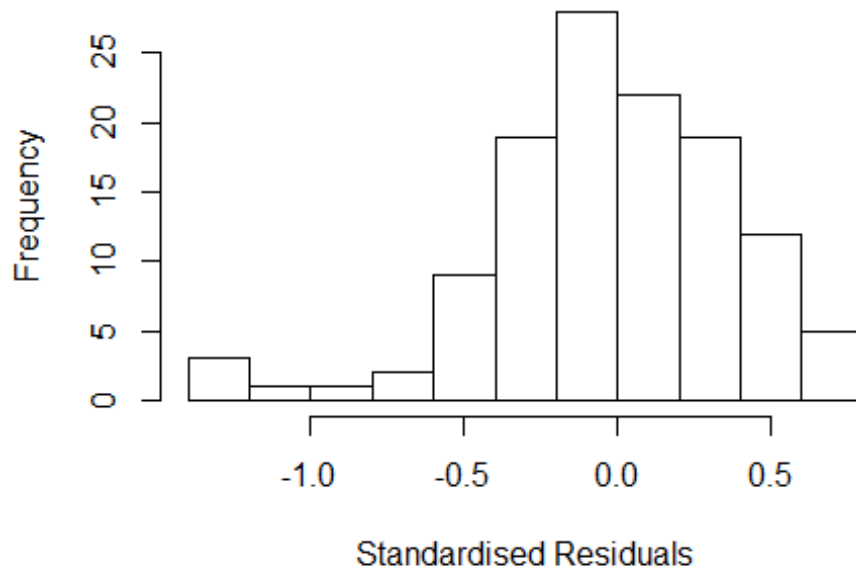
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3304 -0.2692 -0.0122 0.2367 0.7525
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2121 0.1987 6.10 1.9e-08 ***
## FirstAuthorFemale1 0.2543 0.1285 1.98 0.050 .
## LastAuthorFemale1 -0.1000 0.1422 -0.70 0.483
## Year1997 -0.1492 0.2588 -0.58 0.565
## Year1998 -0.0340 0.2872 -0.12 0.906
## Year1999 -0.2156 0.2565 -0.84 0.403
## Year2000 -0.0729 0.2788 -0.26 0.794
## Year2001 -0.0532 0.2533 -0.21 0.834
## Year2002 0.0341 0.2395 0.14 0.887
## Year2003 0.0961 0.3064 0.31 0.754
## Year2004 -0.1255 0.2583 -0.49 0.628
## Year2005 -0.2288 0.2430 -0.94 0.349
```

```

## Year2006          -0.2142      0.2494   -0.86    0.392
## Year2007          -0.2331      0.2351   -0.99    0.324
## Year2008          -0.2572      0.2191   -1.17    0.243
## Year2009          -0.5194      0.2659   -1.95    0.054 .
## Year2010           0.0124      0.2129    0.06    0.954
## Year2011          -0.0265      0.2348   -0.11    0.910
## Year2012           0.1183      0.2279    0.52    0.605
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.391
## Multiple R-squared:  0.176, Adjusted R-squared:  0.0307
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.224  0.882  0.952  0.906  0.990  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      8.26e-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 2.181 1      1.477
## Year              2.181 16      1.025

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3830 -0.2717 -0.0148 0.2367 0.7522
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21194 0.19650 6.17 1.4e-08 ***
## FirstAuthorFemale1 0.25040 0.12455 2.01 0.047 *
## Year1997 -0.13748 0.26265 -0.52 0.602
## Year1998 -0.05288 0.26715 -0.20 0.843
## Year1999 -0.21514 0.25228 -0.85 0.396
## Year2000 -0.07932 0.29612 -0.27 0.789
## Year2001 -0.06724 0.24765 -0.27 0.787
## Year2002 0.03560 0.23298 0.15 0.879
## Year2003 0.09762 0.29839 0.33 0.744
## Year2004 -0.14614 0.25154 -0.58 0.563
## Year2005 -0.22863 0.24063 -0.95 0.344
## Year2006 -0.21415 0.24833 -0.86 0.390
```

```

## Year2007          -0.25195    0.22618   -1.11    0.268
## Year2008          -0.26918    0.21100   -1.28    0.205
## Year2009          -0.54523    0.24572   -2.22    0.029 *
## Year2010          -0.00527    0.21155   -0.02    0.980
## Year2011          -0.03654    0.22696   -0.16    0.872
## Year2012           0.09631    0.22741    0.42    0.673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.177, Adjusted R-squared:  0.0414
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 113 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.192  0.889  0.956  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      8.26e-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 1.635 1      1.278
## Year              1.635 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 ~ 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 -0.2943 -0.0146  0.2686  0.8856

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21161    0.19677    6.16 1.4e-08 ***
## LastAuthorFemale1 -0.08817    0.16029   -0.55    0.58
## Year1997        -0.16422    0.26847   -0.61    0.54
## Year1998         0.00148    0.32924    0.00    1.00
## Year1999        -0.21420    0.25478   -0.84    0.40
## Year2000        -0.06873    0.28660   -0.24    0.81
## Year2001        -0.01221    0.24145   -0.05    0.96
## Year2002         0.11549    0.24026    0.48    0.63
## Year2003         0.15414    0.28474    0.54    0.59
## Year2004        -0.12799    0.25633   -0.50    0.62
## Year2005        -0.22819    0.24123   -0.95    0.35
## Year2006        -0.21401    0.24773   -0.86    0.39
## Year2007        -0.23436    0.23264   -1.01    0.32
## Year2008        -0.19311    0.21427   -0.90    0.37
## Year2009        -0.45826    0.28003   -1.64    0.10
## Year2010         0.03150    0.21676    0.15    0.88
## Year2011         0.00645    0.22776    0.03    0.98
## Year2012         0.20583    0.21114    0.97    0.33
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.399
## Multiple R-squared:  0.143, Adjusted R-squared:  0.00129
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 107 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.180  0.881  0.946  0.905  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      8.26e-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: 121"
## [1] ""

```



```

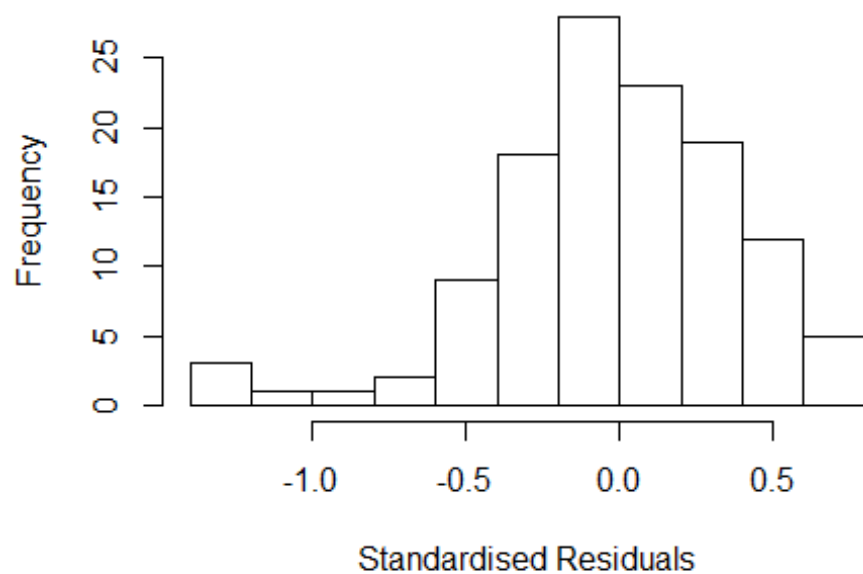
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 1913"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    7    2    1    7    6    3    5    3    2    6    5    8    5
## 2011 2012
##    1    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    5    1    0    2    1    1    3    1    1    2    3    6    3
## 2011 2012
##    0    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    4    1    0    2    0    1    3    0    1    2    3    6    3
## 2011 2012
##    0    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.16227766016838"
## [1] "Male first author team size 2018 geometric mean: 6"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 3.91486764116886"
## [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 lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## Warning in cov2cor(v): diag(.) had 0 or NA entries; non-finite result is
## doubtful

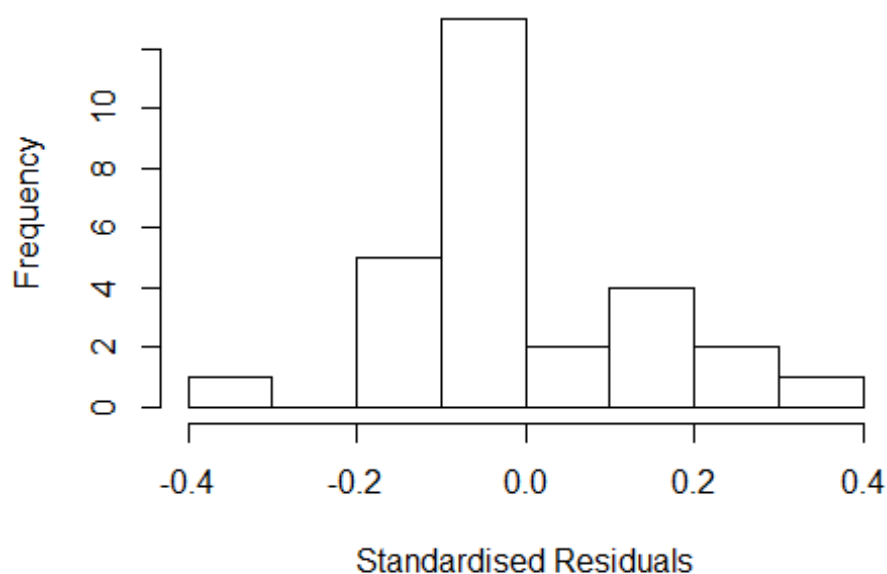
```

Residuals from last author



##		GVIF	Df	$GVIF^{(1/(2*Df))}$
##	FirstAuthorFemale	NaN	1	NaN
##	LastAuthorFemale	NaN	1	NaN
##	Year	NaN	11	NaN

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -0.31435 -0.09358 -0.00353 0.05848 0.37965
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.59e+00 2.75e-09 5.80e+08 < 2e-16 ***
## FirstAuthorFemale1 5.06e-02 5.61e-02 9.00e-01 0.38217
## LastAuthorFemale1 -7.80e-02 8.31e-02 -9.40e-01 0.36390
## Year1998 -4.76e-01 5.80e-02 -8.20e+00 1.0e-06 ***
## Year1999 -3.84e-01 4.53e-09 -8.47e+07 < 2e-16 ***
## Year2001 -1.15e-01 1.32e-01 -8.70e-01 0.39864
## Year2003 1.12e-01 0.00e+00 Inf < 2e-16 ***
## Year2004 -4.57e-01 1.03e-01 -4.42e+00 0.00058 ***
## Year2006 -6.94e-01 0.00e+00 -Inf < 2e-16 ***
## Year2007 -4.04e-01 4.11e-02 -9.83e+00 1.2e-07 ***
## Year2008 -1.75e-01 9.08e-02 -1.92e+00 0.07492 .
## Year2009 -1.13e-01 5.35e-02 -2.11e+00 0.05331 .
```

```

## Year2010          -6.43e-01    2.08e-01 -3.08e+00  0.00811 **
## Year2012          -1.52e-01    0.00e+00      -Inf    < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.242
## Multiple R-squared:  0.682, Adjusted R-squared:  0.386
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 22 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.788  0.957  0.976  0.961  0.996  0.999
## Algorithmic parameters:
##           tuning.chi             bb           tuning.psi           refine.tol
##           1.55e+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-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 -3529  1             NaN
## Year              -3529 11             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
## -3.12e-01 -7.97e-02 -9.44e-16  6.93e-02  3.82e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.59e+00  5.64e-09  2.82e+08 < 2e-16 ***

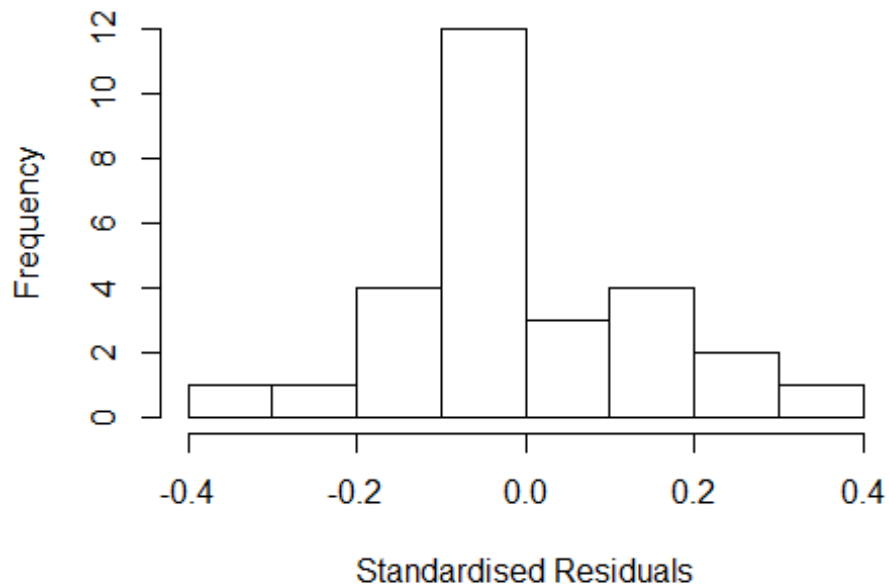
```

```

## FirstAuthorFemale1  4.78e-02  5.79e-02  8.20e-01  0.42246
## Year1998            -4.75e-01  5.81e-02 -8.18e+00  6.5e-07 ***
## Year1999            -3.84e-01  7.57e-09 -5.08e+07  < 2e-16 ***
## Year2001            -1.13e-01  1.35e-01 -8.40e-01  0.41347
## Year2003             1.12e-01  8.92e-10  1.26e+08  < 2e-16 ***
## Year2004            -4.84e-01  9.54e-02 -5.08e+00  0.00014 ***
## Year2006            -6.94e-01  6.73e-09 -1.03e+08  < 2e-16 ***
## Year2007            -4.02e-01  4.25e-02 -9.48e+00  1.0e-07 ***
## Year2008            -1.73e-01  9.22e-02 -1.88e+00  0.07956 .
## Year2009            -1.49e-01  5.33e-02 -2.80e+00  0.01359 *
## Year2010            -6.45e-01  2.15e-01 -3.00e+00  0.00906 **
## Year2012            -1.52e-01  7.62e-09 -2.00e+07  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.229
## Multiple R-squared:  0.677, Adjusted R-squared:  0.419
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 22 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.763  0.952  0.971  0.956  0.993  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      3.57e-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
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

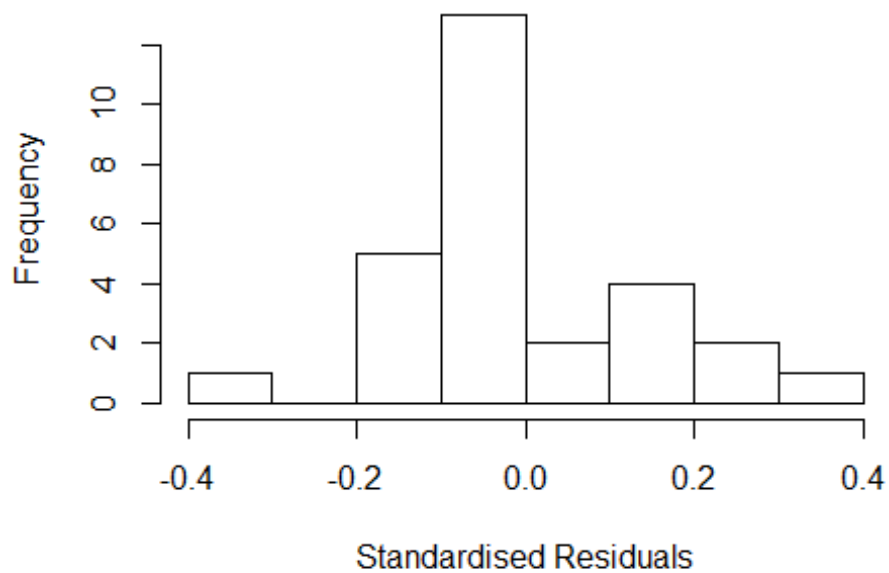
```

Residuals from first author



##	GVIF	Df	$GVIF^{1/(2 \cdot Df)}$
## LastAuthorFemale	NaN	1	NaN
## Year	NaN	11	NaN

Residuals from last author



```

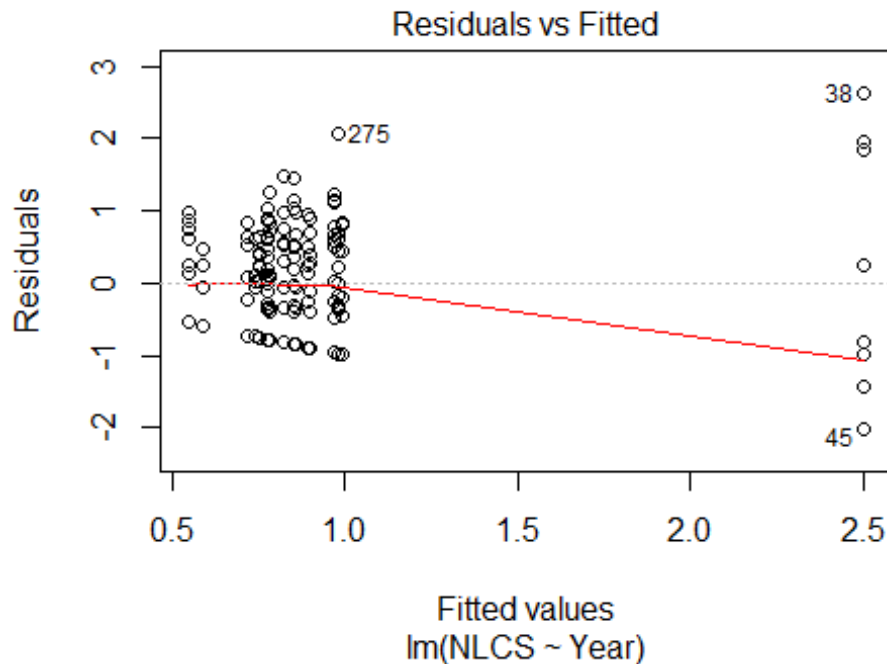
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -3.12e-01 -7.83e-02 -1.39e-17 8.44e-02 3.82e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.59e+00 0.00e+00 Inf < 2e-16 ***
## LastAuthorFemale1 -7.37e-02 7.76e-02 -9.50e-01 0.35745
## Year1998 -4.63e-01 5.06e-02 -9.15e+00 1.6e-07 ***
## Year1999 -3.84e-01 0.00e+00 -Inf < 2e-16 ***
## Year2001 -8.95e-02 1.57e-01 -5.70e-01 0.57663
## Year2003 1.12e-01 0.00e+00 Inf < 2e-16 ***
## Year2004 -4.40e-01 9.76e-02 -4.51e+00 0.00042 ***
## Year2006 -6.94e-01 3.53e-09 -1.96e+08 < 2e-16 ***
## Year2007 -3.78e-01 4.85e-02 -7.80e+00 1.2e-06 ***
## Year2008 -1.44e-01 1.03e-01 -1.39e+00 0.18544
## Year2009 -9.00e-02 5.27e-02 -1.71e+00 0.10834
## Year2010 -6.45e-01 2.18e-01 -2.96e+00 0.00967 **
## Year2012 -1.52e-01 3.53e-09 -4.30e+07 < 2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.225
## Multiple R-squared: 0.678, Adjusted R-squared: 0.42
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 22 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.754 0.947 0.971 0.954 0.992 0.999
## Algorithmic parameters:
## tuning.chi bb tuning.psi refine.tol
## 1.55e+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-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: 28"
## [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
##      8    6    7   11   11   11   13    7   13   11   14   17   16   21   22
## 2011 2012
##    30   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      7    5    7   11    9   10   13    7   12    8   11   17   15   17   21
## 2011 2012
##    19   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      7    5    7   10    9    9   12    6   12    6    8   15   13   14   21
## 2011 2012
##    14   16
## [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

```

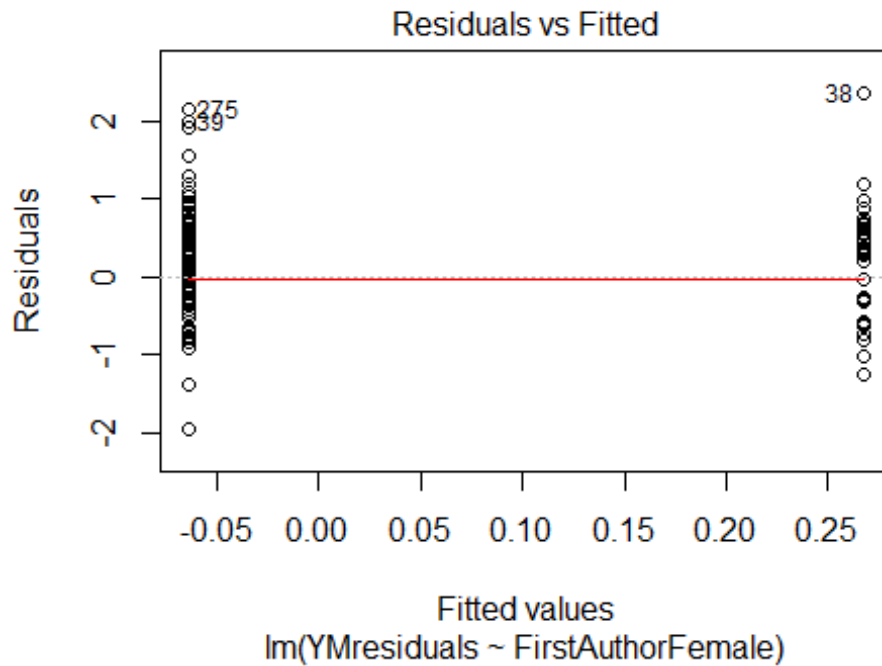
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.5, df = 1, p-value = 0.06

## [1] "Female first author team size 2018 geometric mean: 4"
## [1] "Male first author team size 2018 geometric mean: 1.68731455106541"

## 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 = 19, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.40187391035201"
## [1] "Male last author team size 2018 geometric mean: 1.54258405896365"

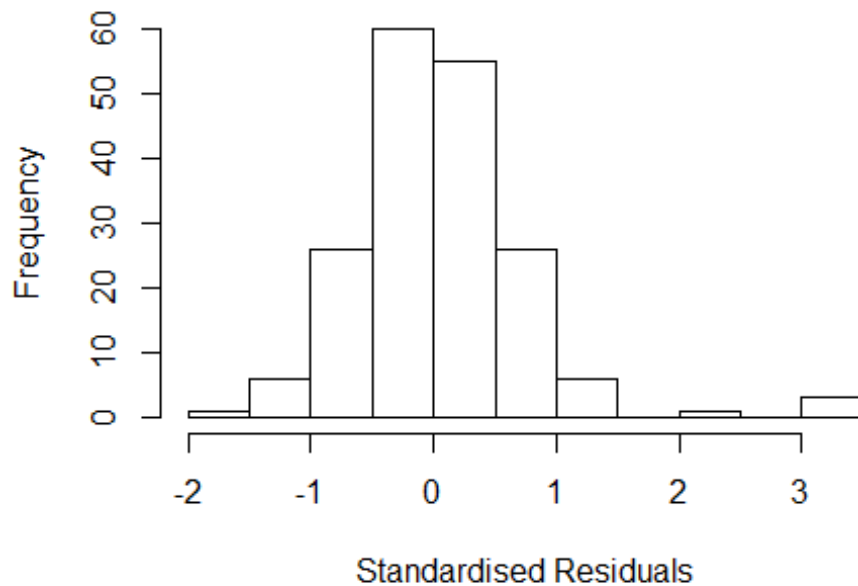
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 62, 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	2.322	1	1.524
LastAuthorFemale	3.210	1	1.792
UniqueAuthors	229.279	4	1.973
Year	431.369	16	1.209

Residuals from first and last author and team size



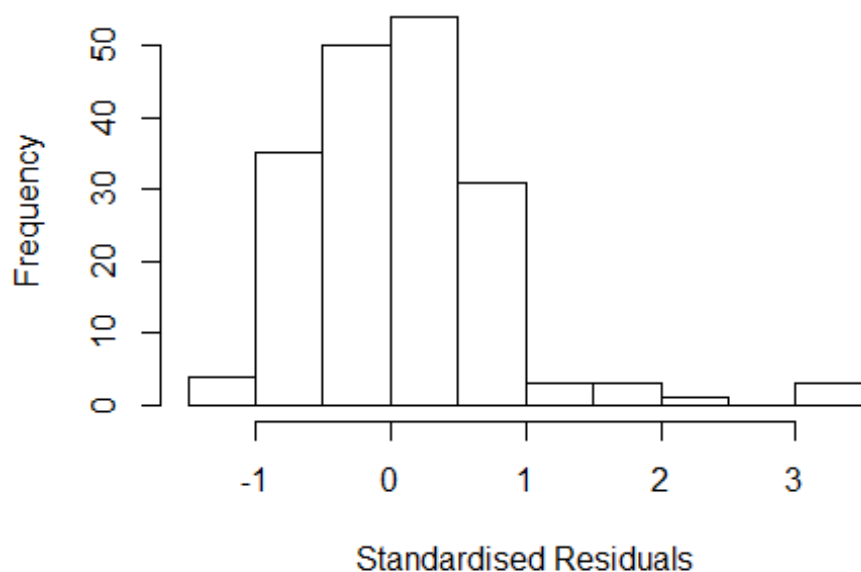
```
## [1] "List of 3 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 36  0009459892 4.357 2000      2000      1      3.196
## 38  84909947725 5.125 2000      2000      1      3.183
## 39  85011136862 4.445 2000      2000      1      3.284
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.756119 -0.377076 -0.000463  0.420996  3.283610
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8861    0.2628   3.37 0.00093 ***
## FirstAuthorFemale1 0.2488    0.1923   1.29 0.19756
## LastAuthorFemale1 0.3787    0.1807   2.10 0.03771 *
## UniqueAuthors2    0.1526    0.1510   1.01 0.31351
## UniqueAuthors3    0.4885    0.2541   1.92 0.05633 .
## UniqueAuthors4    0.8962    0.1689   5.30 3.7e-07 ***
## UniqueAuthors5    0.0658    0.2043   0.32 0.74782
## Year1997        -0.2903    0.3108  -0.93 0.35168
```

```

## Year1998          -0.0665      0.3482    -0.19  0.84890
## Year1999          -0.0766      0.3237    -0.24  0.81326
## Year2000           0.2753      0.3888     0.71  0.47989
## Year2001          -0.3982      0.3057    -1.30  0.19467
## Year2002          -0.1669      0.3498    -0.48  0.63399
## Year2003          -0.2894      0.3069    -0.94  0.34709
## Year2004          -0.4255      0.2819    -1.51  0.13309
## Year2005          -0.2850      0.4330    -0.66  0.51130
## Year2006          -0.1327      0.3043    -0.44  0.66342
## Year2007          -0.5327      0.3003    -1.77  0.07800 .
## Year2008          -0.5039      0.3331    -1.51  0.13230
## Year2009          -0.2920      0.3191    -0.92  0.36149
## Year2010          -0.2907      0.2920    -1.00  0.32100
## Year2011          -0.2033      0.3416    -0.60  0.55255
## Year2012          -0.2460      0.3724    -0.66  0.50980
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.558
## Multiple R-squared:  0.211, Adjusted R-squared:  0.103
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 3 observations c(31,32,33) are outliers with |weight| = 0 ( < 0.00054);
## 13 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0209 0.8720 0.9530 0.9000 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          5.43e-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 2.202 1          1.484
## LastAuthorFemale 2.081 1          1.442
## Year 2.427 16          1.028

```

Residuals from first and last author



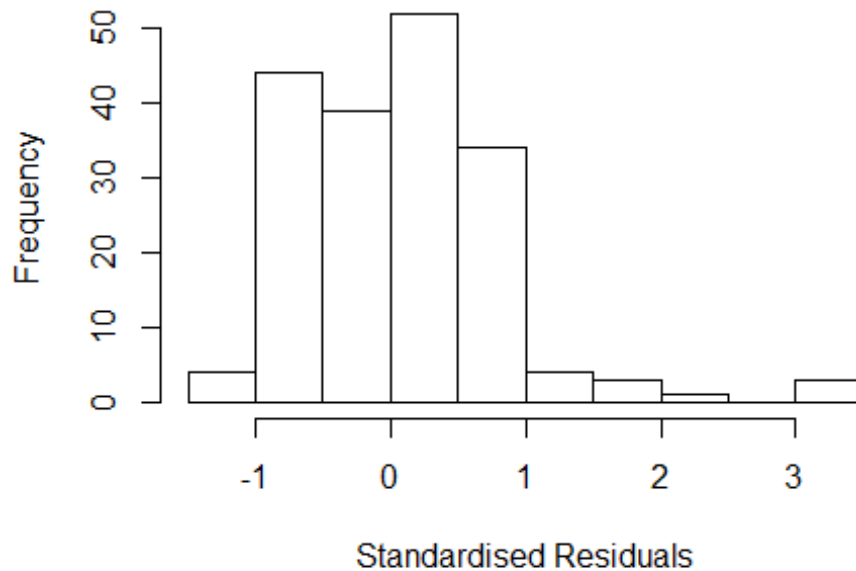
```
## [1] "List of 3 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 36  0009459892 4.357 2000      2000      1      3.154
## 38  84909947725 5.125 2000      2000      1      3.308
## 39  85011136862 4.445 2000      2000      1      3.242
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4868 -0.4547  0.0322  0.4293  3.3078
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8979    0.2437   3.68 0.00031 ***
## FirstAuthorFemale1  0.2376    0.2043   1.16 0.24663
## LastAuthorFemale1  0.3768    0.1716   2.20 0.02949 *
## Year1997        -0.3027    0.2944  -1.03 0.30524
## Year1998        -0.0648    0.3414  -0.19 0.84972
## Year1999        -0.0200    0.3156  -0.06 0.94950
## Year2000         0.3050    0.3740   0.82 0.41600
## Year2001        -0.3719    0.2877  -1.29 0.19804
## Year2002        -0.1397    0.3010  -0.46 0.64309
## Year2003        -0.2632    0.2786  -0.94 0.34628
```

```

## Year2004          -0.3183      0.2596    -1.23   0.22184
## Year2005          -0.1833      0.3625    -0.51   0.61386
## Year2006          -0.0913      0.3057    -0.30   0.76549
## Year2007          -0.3457      0.3036    -1.14   0.25656
## Year2008          -0.4166      0.3338    -1.25   0.21383
## Year2009          -0.2284      0.3108    -0.73   0.46342
## Year2010          -0.2340      0.2609    -0.90   0.37109
## Year2011          -0.1665      0.3239    -0.51   0.60797
## Year2012          -0.0255      0.3129    -0.08   0.93518
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.595
## Multiple R-squared:  0.147, Adjusted R-squared:  0.0541
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 3 observations c(31,32,33) are outliers with |weight| = 0 ( < 0.00054);
## 11 weights are ~= 1. The remaining 170 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.147  0.873   0.943   0.907   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      5.43e-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 1.677 1      1.295
## Year              1.677 16      1.016

```

Residuals from first author



```
## [1] "List of 3 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 36  0009459892 4.357 2000      2000      1      3.154
## 38  84909947725 5.125 2000      2000      1      3.308
## 39  85011136862 4.445 2000      2000      1      3.242
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3275 -0.5086  0.0301  0.4836  3.4230
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9245    0.2510     3.68  0.00031 ***
## FirstAuthorFemale1 0.4589    0.1772     2.59  0.01049 *
## Year1997       -0.3296    0.3002    -1.10  0.27392
## Year1998       -0.0908    0.3467    -0.26  0.79379
## Year1999       -0.0468    0.3209    -0.15  0.88416
## Year2000        0.3186    0.3774     0.84  0.39979
## Year2001       -0.3195    0.3143    -1.02  0.31084
## Year2002       -0.1686    0.3014    -0.56  0.57651
## Year2003       -0.2756    0.3033    -0.91  0.36481
## Year2004       -0.2948    0.2659    -1.11  0.26910
```

```

## Year2005          -0.1701      0.3727   -0.46  0.64870
## Year2006          -0.1019      0.3205   -0.32  0.75078
## Year2007          -0.3612      0.3099   -1.17  0.24540
## Year2008          -0.4160      0.3459   -1.20  0.23090
## Year2009          -0.2464      0.3251   -0.76  0.44961
## Year2010          -0.2163      0.2711   -0.80  0.42623
## Year2011          -0.1389      0.3402   -0.41  0.68358
## Year2012          -0.0558      0.3181   -0.18  0.86088
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.602
## Multiple R-squared:  0.116, Adjusted R-squared:  0.0258
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 3 observations c(31,32,33) are outliers with |weight| = 0 ( < 0.00054);
## 17 weights are ~= 1. The remaining 164 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.157  0.863  0.928   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      5.43e-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 1.738 1          1.318
## Year          1.738 16          1.017
##
## [1] "List of 3 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 36  0009459892 4.357 2000      2000      1      3.154
## 38  84909947725 5.125 2000      2000      1      3.308
## 39  85011136862 4.445 2000      2000      1      3.242
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,

```



```

##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.409 -0.452  0.030  0.457  3.383
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9163     0.2470   3.71  0.00028 ***
## LastAuthorFemale1 0.5084     0.1626   3.13  0.00209 **
## Year1997         -0.3214     0.2968  -1.08  0.28047
## Year1998         -0.0822     0.3435  -0.24  0.81119
## Year1999         -0.0387     0.3176  -0.12  0.90321
## Year2000          0.3168     0.3754   0.84  0.39985
## Year2001         -0.3809     0.2945  -1.29  0.19778
## Year2002         -0.1613     0.2992  -0.54  0.59052
## Year2003         -0.2386     0.2725  -0.88  0.38269
## Year2004         -0.3095     0.2722  -1.14  0.25716
## Year2005         -0.2131     0.3765  -0.57  0.57224
## Year2006         -0.0984     0.3146  -0.31  0.75475
## Year2007         -0.3090     0.3082  -1.00  0.31745
## Year2008         -0.4419     0.3260  -1.36  0.17710
## Year2009         -0.1979     0.3199  -0.62  0.53697
## Year2010         -0.2465     0.2684  -0.92  0.35968
## Year2011         -0.1928     0.3251  -0.59  0.55399
## Year2012         -0.0159     0.3092  -0.05  0.95917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.606
## Multiple R-squared:  0.132, Adjusted R-squared:  0.0428
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 3 observations c(31,32,33) are outliers with |weight| = 0 ( < 0.00054);
## 12 weights are ~= 1. The remaining 169 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.178  0.876  0.945  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          5.43e-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

```

```

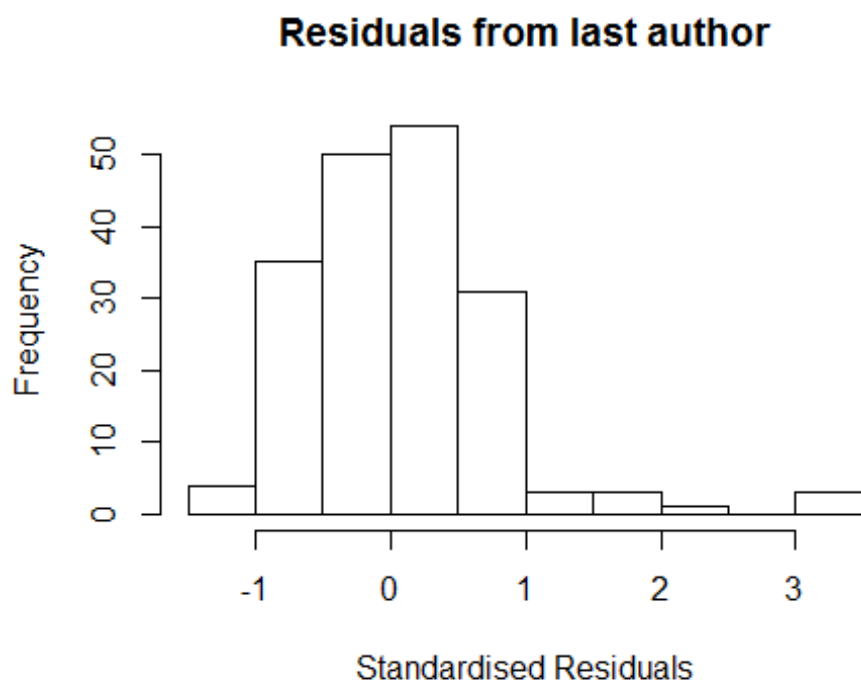
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 184"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2001"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    1    3    8    4    2    4   10    3    8    1    1   10    7   13
## 2011 2012
##   11    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    3    7    1    1    4    9    1    8    1    1   10    6    6
## 2011 2012
##    6    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    3    7    1    1    4    9    1    8    1    1    9    6    6
## 2011 2012
##    6    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1"
## [1] "Male first author team size 2018 geometric mean: 1.96259879303011"

## 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 = 1, p-value = 0.3
## 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: 1.81712059283214"

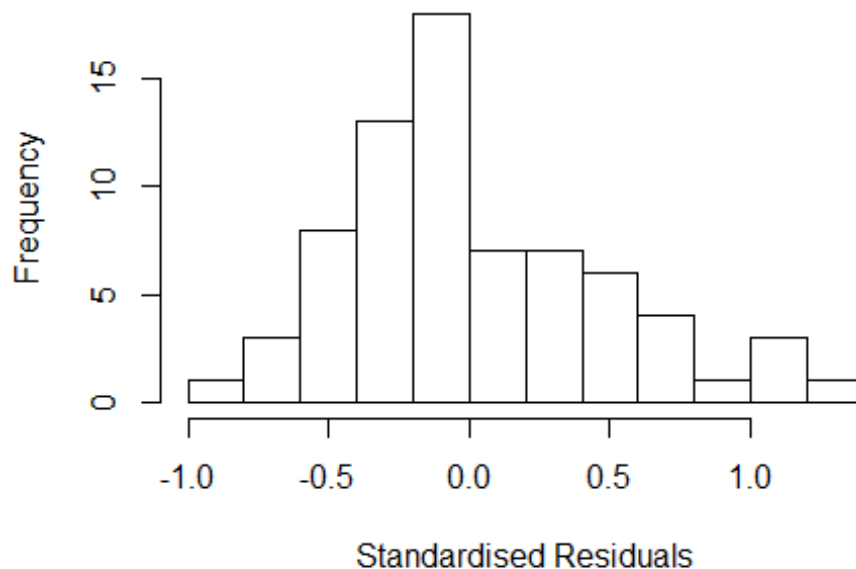
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, 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 -1.304e+15  1          NaN
## LastAuthorFemale  -6.503e+14  1          NaN
## UniqueAuthors      5.362e+58  4      2.194e+07
## Year                4.544e+72 16      1.864e+02
```

Residuals from first and last author and team size



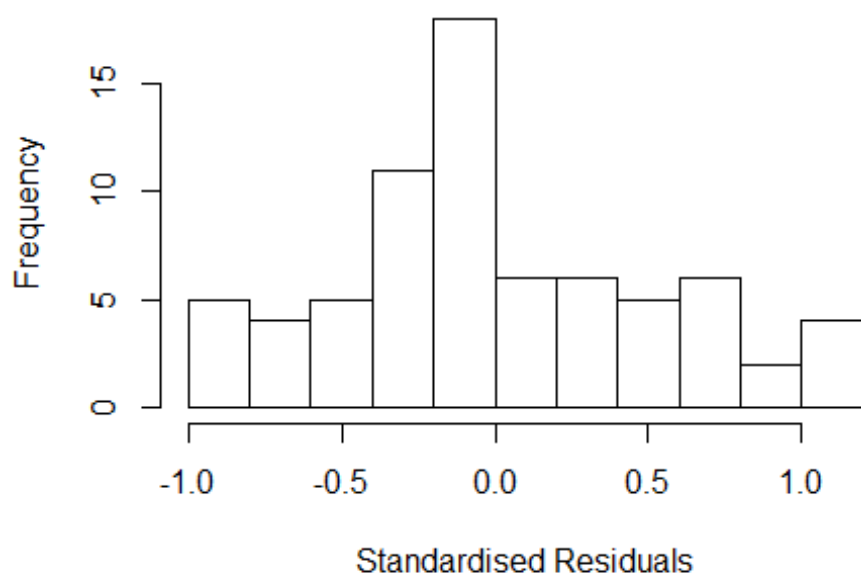
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.95e-01 -3.31e-01 -8.60e-16 3.03e-01 1.28e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.343 0.321 1.07 0.29030
## FirstAuthorFemale1 -0.912 0.179 -5.09 5.7e-06 ***
## LastAuthorFemale1 0.739 0.194 3.80 0.00040 ***
## UniqueAuthors2 0.469 0.184 2.55 0.01409 *
## UniqueAuthors3 0.129 0.207 0.62 0.53490
## UniqueAuthors4 0.983 0.244 4.04 0.00019 ***
## UniqueAuthors5 0.969 0.455 2.13 0.03804 *
## Year1997 1.760 0.318 5.53 1.2e-06 ***
## Year1998 0.161 0.396 0.41 0.68710
## Year1999 0.926 0.391 2.37 0.02194 *
```

```

## Year2000          1.006      0.321      3.13  0.00292 **
## Year2001          0.822      0.263      3.13  0.00298 **
## Year2002          1.067      0.350      3.05  0.00373 **
## Year2003          0.397      0.311      1.28  0.20739
## Year2004         -0.385      0.263     -1.46  0.14957
## Year2005          0.492      0.333      1.48  0.14570
## Year2006          0.129      0.263      0.49  0.62594
## Year2007         -0.355      0.370     -0.96  0.34172
## Year2008          0.223      0.340      0.66  0.51545
## Year2009          0.601      0.362      1.66  0.10313
## Year2010          0.627      0.405      1.55  0.12829
## Year2011          0.336      0.455      0.74  0.46419
## Year2012          0.419      0.456      0.92  0.36200
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.541
## Multiple R-squared:  0.392, Adjusted R-squared:  0.119
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.555  0.895  0.953  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      1.39e-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 2.338e+14 1      1.529e+07
## LastAuthorFemale  2.875e+14 1      1.695e+07
## Year              5.137e+28 16      7.892e+00

```

Residuals from first and last author



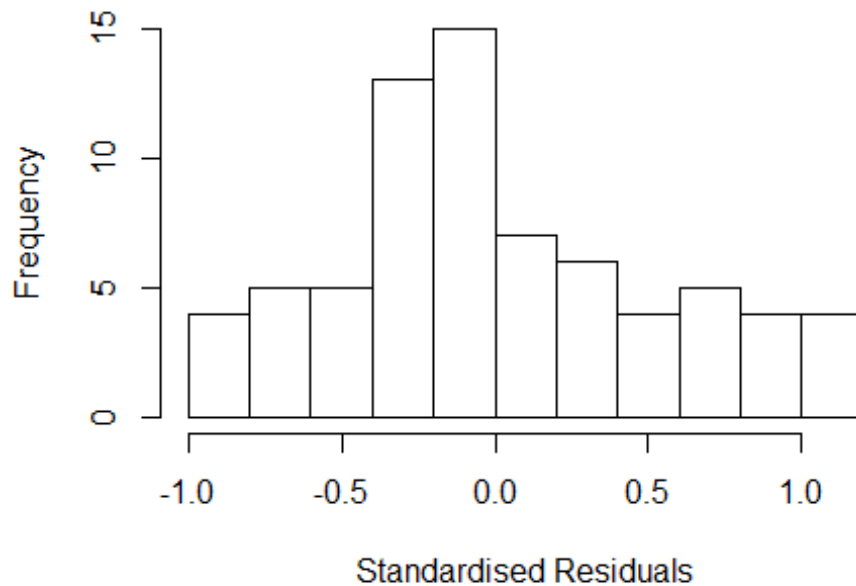
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.88990 -0.35649 -0.00716 0.33411 1.18975
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8120 0.2594 3.13 0.00284 **
## FirstAuthorFemale1 -0.7543 0.2004 -3.76 0.00042 ***
## LastAuthorFemale1 0.5170 0.2156 2.40 0.02005 *
## Year1997 1.6023 0.3277 4.89 9.8e-06 ***
## Year1998 -0.1166 0.4618 -0.25 0.80166
## Year1999 0.5643 0.3516 1.60 0.11446
## Year2000 0.5370 0.2594 2.07 0.04330 *
## Year2001 0.8220 0.2594 3.17 0.00254 **
## Year2002 0.6935 0.3316 2.09 0.04130 *
## Year2003 0.1804 0.3089 0.58 0.56155
## Year2004 -0.3850 0.2594 -1.48 0.14362
## Year2005 0.2482 0.3154 0.79 0.43490
```

```

## Year2006          0.1290      0.2594      0.50  0.62098
## Year2007         -0.4720      0.3373     -1.40  0.16751
## Year2008          0.0779      0.3439      0.23  0.82167
## Year2009          0.3445      0.3699      0.93  0.35588
## Year2010          0.4653      0.3814      1.22  0.22781
## Year2011          0.0250      0.5056      0.05  0.96077
## Year2012          0.2481      0.3889      0.64  0.52617
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.594
## Multiple R-squared:  0.256, Adjusted R-squared:  0.00328
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.668  0.865   0.949   0.918   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      1.39e-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 1.646e+14 1      1.283e+07
## Year              1.646e+14 16      2.781e+00

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.9259 -0.3343 -0.0406  0.3573  1.1714
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.8120     0.2569   3.16 0.00258 **
## FirstAuthorFemale1 -0.3686     0.2018  -1.83 0.07326 .
## Year1997          1.2166     0.3267   3.72 0.00047 ***
## Year1998          -0.0631     0.4298  -0.15 0.88374
## Year1999           0.5826     0.3556   1.64 0.10718
## Year2000           0.5370     0.2569   2.09 0.04134 *
## Year2001           0.8220     0.2569   3.20 0.00231 **
## Year2002           0.6291     0.3083   2.04 0.04620 *
## Year2003           0.1500     0.3129   0.48 0.63359
## Year2004          -0.3850     0.2569  -1.50 0.13984
## Year2005           0.3338     0.3123   1.07 0.28986
## Year2006           0.1290     0.2569   0.50 0.61765
```

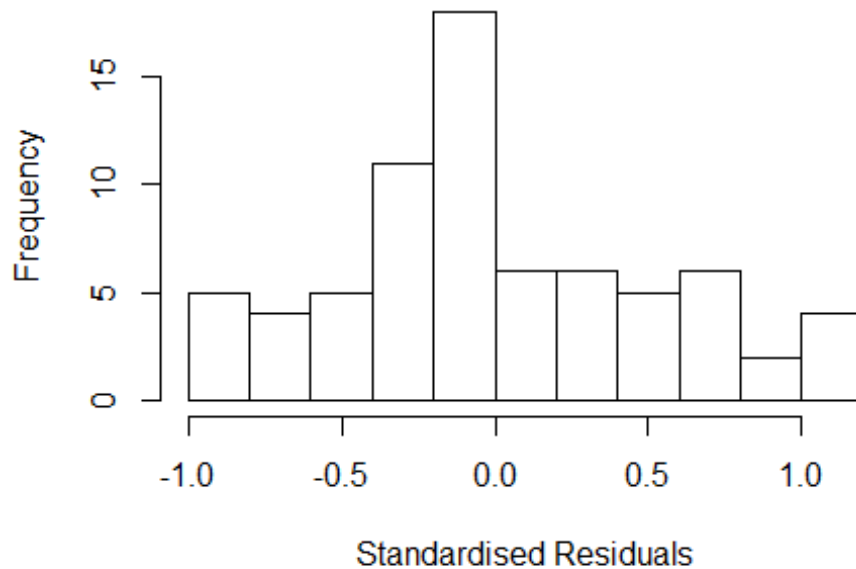


```

## Year2007          0.0450      0.2569      0.18  0.86162
## Year2008          0.1139      0.3476      0.33  0.74444
## Year2009          0.4512      0.4259      1.06  0.29412
## Year2010          0.6135      0.3812      1.61  0.11331
## Year2011          0.1024      0.4637      0.22  0.82596
## Year2012          0.2066      0.3527      0.59  0.56058
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.21,   Adjusted R-squared:  -0.0386
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 63 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.718  0.889  0.966  0.929  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-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 -6.02e+14  1          NaN
## Year             -6.02e+14 16          NaN

```

Residuals from last author



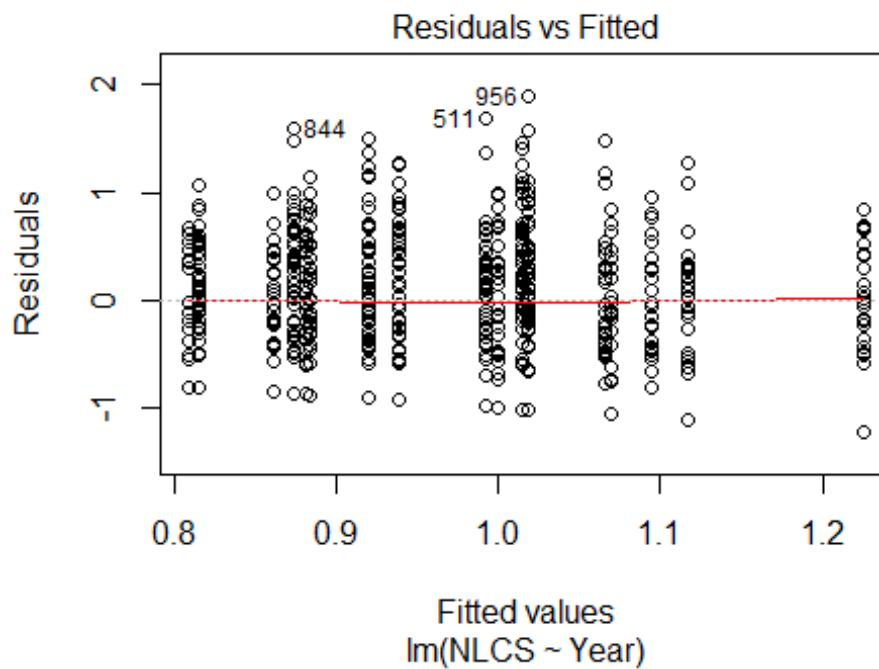
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.071 -0.346 -0.017 0.380 1.376
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8120 0.2582 3.15 0.0027 **
## LastAuthorFemale1 -0.0621 0.2327 -0.27 0.7907
## Year1997 0.8480 0.2582 3.28 0.0018 **
## Year1998 -0.1755 0.4953 -0.35 0.7244
## Year1999 0.3780 0.3904 0.97 0.3373
## Year2000 0.5370 0.2582 2.08 0.0423 *
## Year2001 0.8220 0.2582 3.18 0.0024 **
## Year2002 0.4539 0.3103 1.46 0.1492
## Year2003 0.0670 0.3020 0.22 0.8253
## Year2004 -0.3850 0.2582 -1.49 0.1417
## Year2005 0.3212 0.3272 0.98 0.3306
## Year2006 0.1290 0.2582 0.50 0.6193
```

```

## Year2007          0.1071      0.3476      0.31      0.7593
## Year2008          0.0358      0.3352      0.11      0.9154
## Year2009          0.3155      0.4013      0.79      0.4353
## Year2010          0.5133      0.4044      1.27      0.2097
## Year2011         -0.0758      0.5234     -0.14      0.8854
## Year2012          0.0442      0.3919      0.11      0.9105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.148, Adjusted R-squared:  -0.12
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.596  0.870   0.933   0.914   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      1.39e-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: 72"
## [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
##   35   35   36   35   40   31   43   32   52   34   59   67   73   78   82
## 2011 2012
##   72   74
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   21   26   29   26   32   17   36   28   39   31   50   57   55   59   66
## 2011 2012

```

```
## 54 59
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 20 23 29 23 31 15 34 26 35 30 46 53 50 55 59
## 2011 2012
## 46 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 = 24, df = 16, p-value = 0.09
```



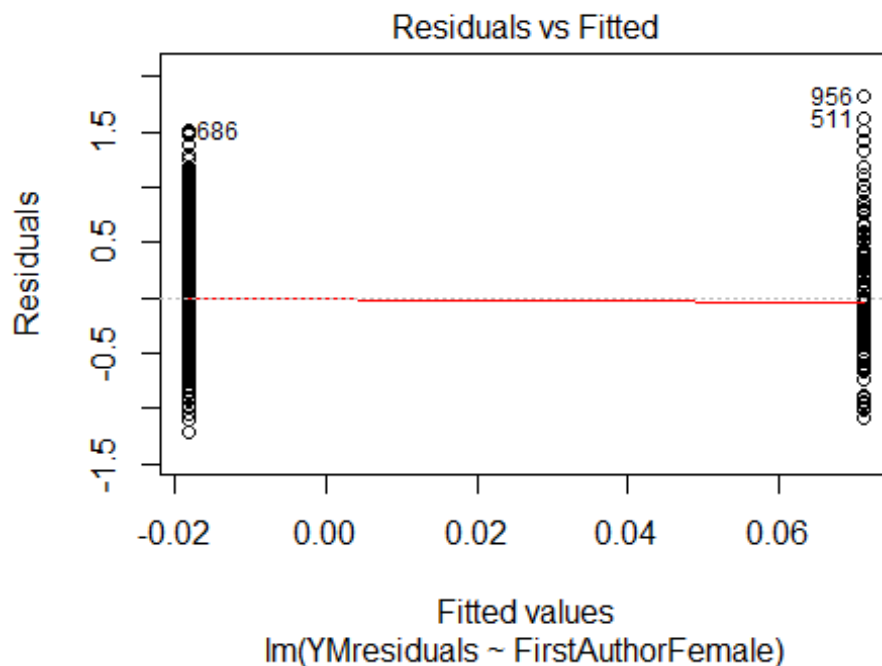
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.5, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 2.14605459236625"
## [1] "Male first author team size 2018 geometric mean: 1.83573822595812"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

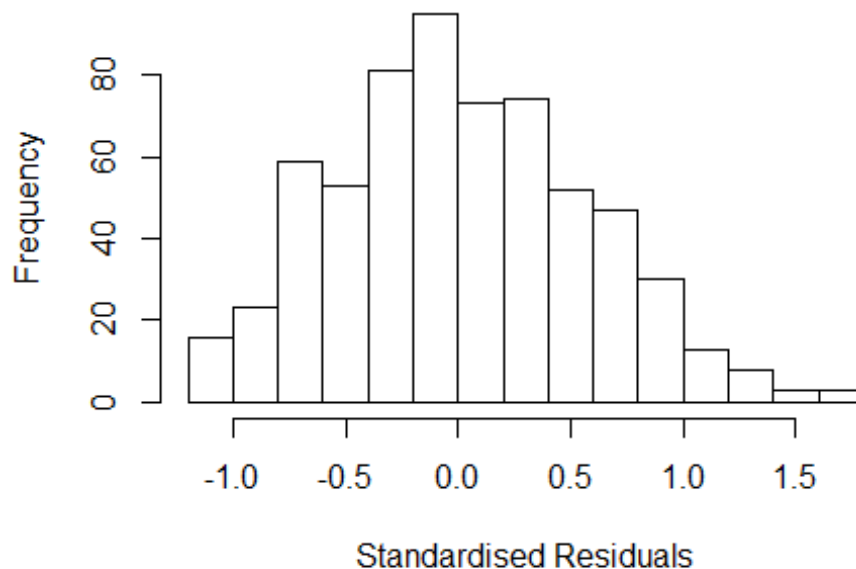
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 340, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.11139358420235"
## [1] "Male last author team size 2018 geometric mean: 1.83873526114933"

## 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 = 350, 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 1.855 1      1.362
## LastAuthorFemale  1.934 1      1.391
## UniqueAuthors    1.828 4      1.078
## Year              2.176 16     1.025
```

Residuals from first and last author and team size



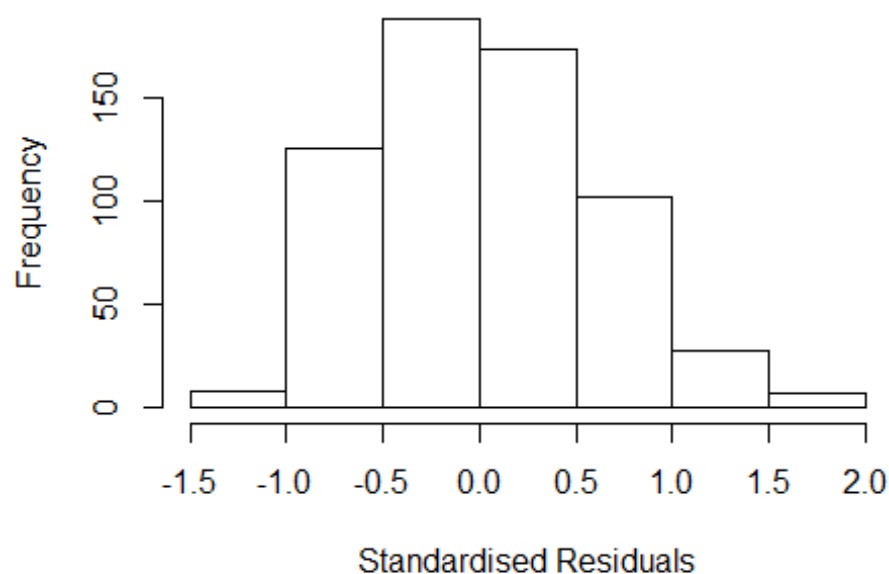
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1620 -0.3859 -0.0213 0.3950 1.7660
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0216 0.1041 9.82 < 2e-16 ***
## FirstAuthorFemale1 0.0223 0.0804 0.28 0.78197
## LastAuthorFemale1 0.0274 0.0793 0.35 0.72940
## UniqueAuthors2 0.2453 0.0566 4.33 1.7e-05 ***
## UniqueAuthors3 0.3034 0.0872 3.48 0.00054 ***
## UniqueAuthors4 0.5780 0.1779 3.25 0.00122 **
## UniqueAuthors5 0.6456 0.1028 6.28 6.5e-10 ***
## Year1997 0.0860 0.1488 0.58 0.56370
## Year1998 -0.0830 0.1346 -0.62 0.53779
## Year1999 0.0309 0.1509 0.20 0.83798
```

```

## Year2000          -0.1274      0.1452   -0.88   0.38083
## Year2001          -0.2562      0.1564   -1.64   0.10188
## Year2002          -0.2223      0.1356   -1.64   0.10177
## Year2003          -0.0435      0.1509   -0.29   0.77308
## Year2004          -0.3187      0.1346   -2.37   0.01824 *
## Year2005          -0.3046      0.1281   -2.38   0.01775 *
## Year2006          -0.1562      0.1249   -1.25   0.21143
## Year2007          -0.3379      0.1218   -2.77   0.00570 **
## Year2008          -0.2611      0.1377   -1.90   0.05843 .
## Year2009          -0.2941      0.1427   -2.06   0.03973 *
## Year2010          -0.1936      0.1279   -1.51   0.13071
## Year2011          -0.2646      0.1422   -1.86   0.06321 .
## Year2012          -0.2127      0.1534   -1.39   0.16615
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.588
## Multiple R-squared:  0.0936, Adjusted R-squared:  0.0608
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 58 weights are ~= 1. The remaining 572 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.347  0.873  0.953   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          1.59e-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 1.876 1          1.370
## LastAuthorFemale  1.884 1          1.373
## Year              1.205 16          1.006

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.20691 -0.42770 -0.00765 0.41259 1.93007
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1111 0.1141 9.74 <2e-16 ***
## FirstAuthorFemale1 0.0335 0.0837 0.40 0.689
## LastAuthorFemale1 0.0526 0.0824 0.64 0.523
## Year1997 0.0958 0.1565 0.61 0.541
## Year1998 -0.0959 0.1498 -0.64 0.522
## Year1999 0.0226 0.1549 0.15 0.884
## Year2000 -0.1428 0.1577 -0.91 0.366
## Year2001 -0.3020 0.1631 -1.85 0.065 .
## Year2002 -0.2814 0.1417 -1.99 0.048 *
## Year2003 -0.0444 0.1578 -0.28 0.779
## Year2004 -0.2917 0.1460 -2.00 0.046 *
## Year2005 -0.2715 0.1437 -1.89 0.059 .
```

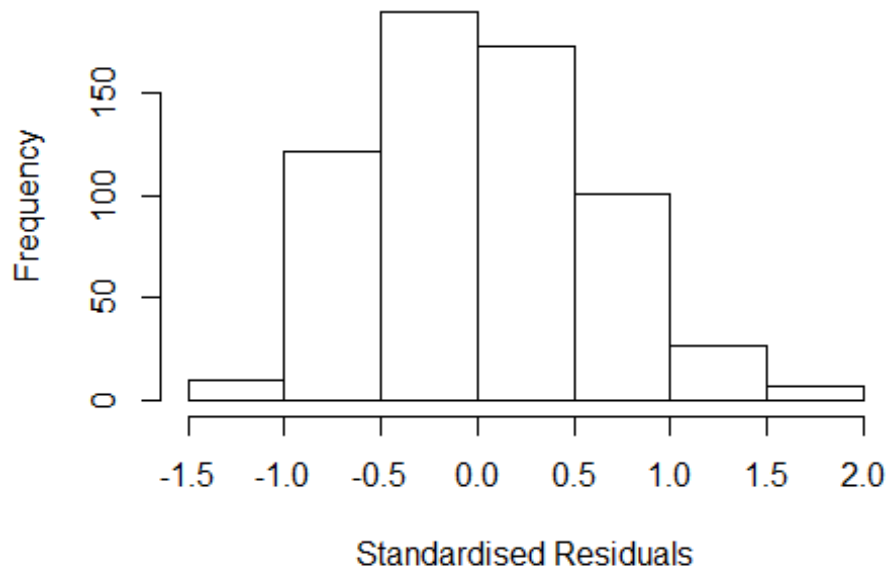


```

## Year2006          -0.1002      0.1359   -0.74    0.461
## Year2007          -0.3030      0.1320   -2.30    0.022 *
## Year2008          -0.2649      0.1509   -1.76    0.080 .
## Year2009          -0.2701      0.1515   -1.78    0.075 .
## Year2010          -0.1637      0.1402   -1.17    0.243
## Year2011          -0.1989      0.1535   -1.30    0.196
## Year2012          -0.1557      0.1614   -0.96    0.335
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.614
## Multiple R-squared:  0.0338, Adjusted R-squared:  0.0053
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 49 weights are ~= 1. The remaining 581 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.302  0.882  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      1.59e-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 1.109 1          1.053
## Year              1.109 16          1.003

```

Residuals from first author



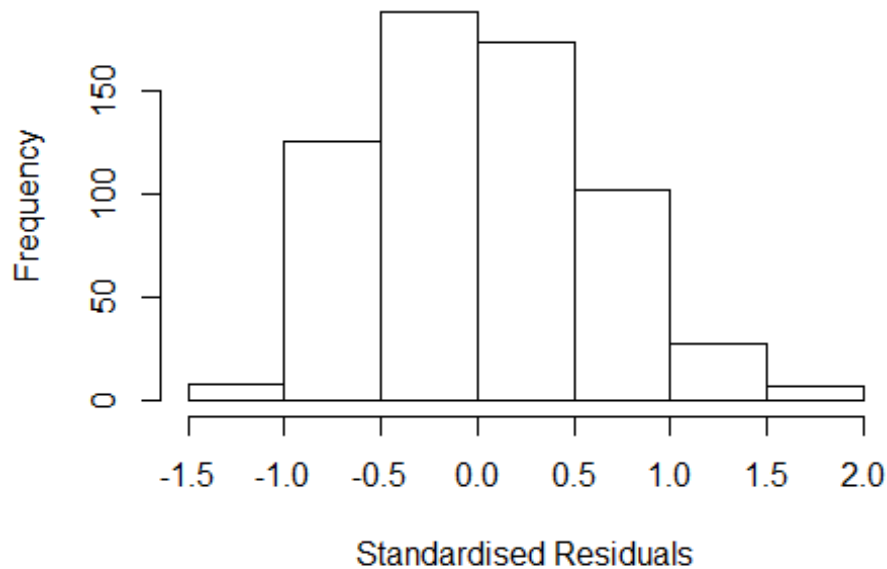
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20687 -0.42887 -0.00725 0.40901 1.89879
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1147 0.1127 9.89 <2e-16 ***
## FirstAuthorFemale1 0.0650 0.0641 1.01 0.311
## Year1997 0.0922 0.1552 0.59 0.553
## Year1998 -0.0999 0.1492 -0.67 0.504
## Year1999 0.0184 0.1537 0.12 0.904
## Year2000 -0.1423 0.1569 -0.91 0.365
## Year2001 -0.3068 0.1629 -1.88 0.060 .
## Year2002 -0.2838 0.1407 -2.02 0.044 *
## Year2003 -0.0433 0.1563 -0.28 0.782
## Year2004 -0.2889 0.1449 -1.99 0.047 *
## Year2005 -0.2622 0.1426 -1.84 0.066 .
## Year2006 -0.0982 0.1352 -0.73 0.468
```

```

## Year2007          -0.2975      0.1309    -2.27      0.023 *
## Year2008          -0.2610      0.1510    -1.73      0.084 .
## Year2009          -0.2690      0.1504    -1.79      0.074 .
## Year2010          -0.1675      0.1390    -1.21      0.229
## Year2011          -0.1948      0.1522    -1.28      0.201
## Year2012          -0.1524      0.1605    -0.95      0.343
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.613
## Multiple R-squared:  0.0331, Adjusted R-squared:  0.00626
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 53 weights are ~= 1. The remaining 577 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.317  0.882  0.948  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      1.59e-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 1.105 1      1.051
## Year              1.105 16      1.003

```

Residuals from last author



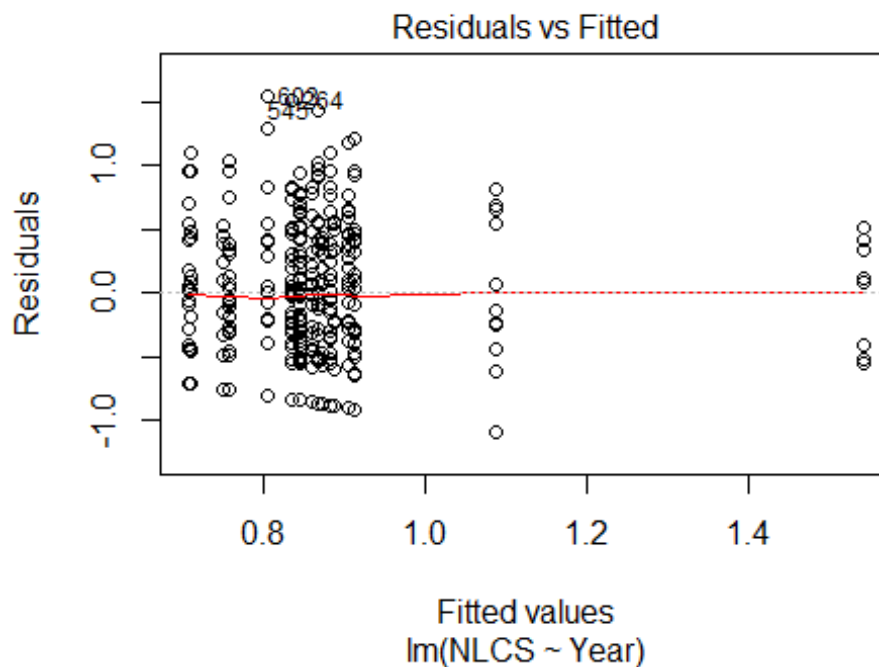
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2117 -0.4268 -0.0104 0.4029 1.9599
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1134 0.1142 9.75 <2e-16 ***
## LastAuthorFemale1 0.0720 0.0630 1.14 0.254
## Year1997 0.0983 0.1573 0.63 0.532
## Year1998 -0.0965 0.1498 -0.64 0.519
## Year1999 0.0227 0.1554 0.15 0.884
## Year2000 -0.1452 0.1578 -0.92 0.358
## Year2001 -0.3029 0.1625 -1.86 0.063 .
## Year2002 -0.2796 0.1424 -1.96 0.050 .
## Year2003 -0.0430 0.1580 -0.27 0.786
## Year2004 -0.2932 0.1461 -2.01 0.045 *
## Year2005 -0.2741 0.1432 -1.91 0.056 .
## Year2006 -0.1010 0.1361 -0.74 0.458
```

```

## Year2007          -0.3045      0.1319    -2.31      0.021 *
## Year2008          -0.2653      0.1510    -1.76      0.079 .
## Year2009          -0.2699      0.1516    -1.78      0.075 .
## Year2010          -0.1623      0.1403    -1.16      0.248
## Year2011          -0.1959      0.1538    -1.27      0.203
## Year2012          -0.1547      0.1617    -0.96      0.339
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.614
## Multiple R-squared:  0.0335, Adjusted R-squared:  0.00661
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 50 weights are ~= 1. The remaining 580 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.287  0.882  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      1.59e-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: 630"
## [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
##   14   10   17   11   16   15   11   18   18   26   36   30   41   36   41
## 2011 2012
##   38   30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11    8   14    9   13   11    8   18   18   24   33   26   29   28   30
## 2011 2012

```

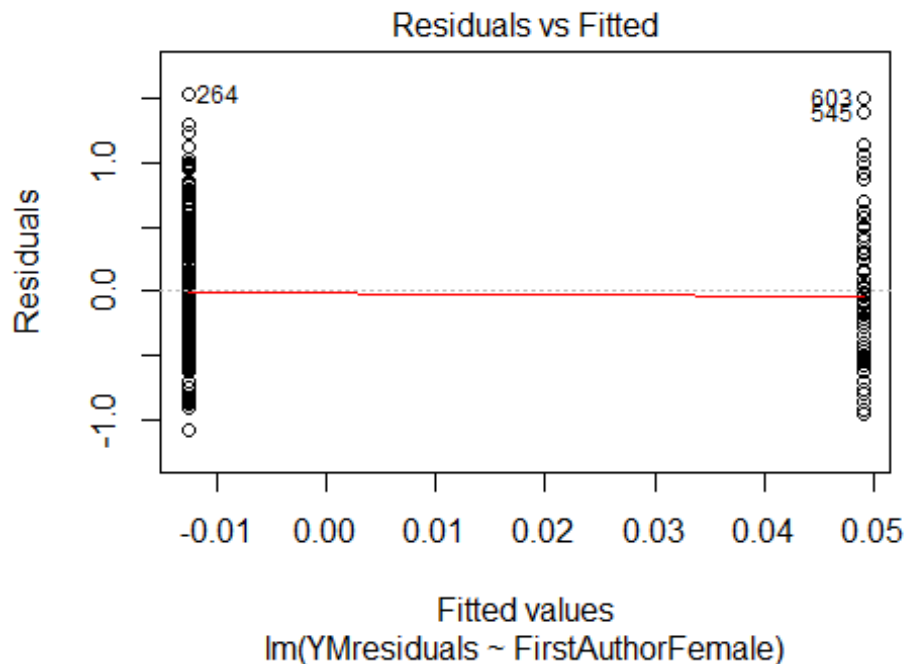
```
## 26 22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 7 14 8 13 11 8 18 16 23 31 23 24 26 27
## 2011 2012
## 24 18
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 12, df = 16, p-value = 0.7
```



```
##
## 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: 1.86606598307361"
## [1] "Male first author team size 2018 geometric mean: 1.76058884973737"
## 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: 1.9185101051426"
## [1] "Male last author team size 2018 geometric mean: 1.7370729389807"

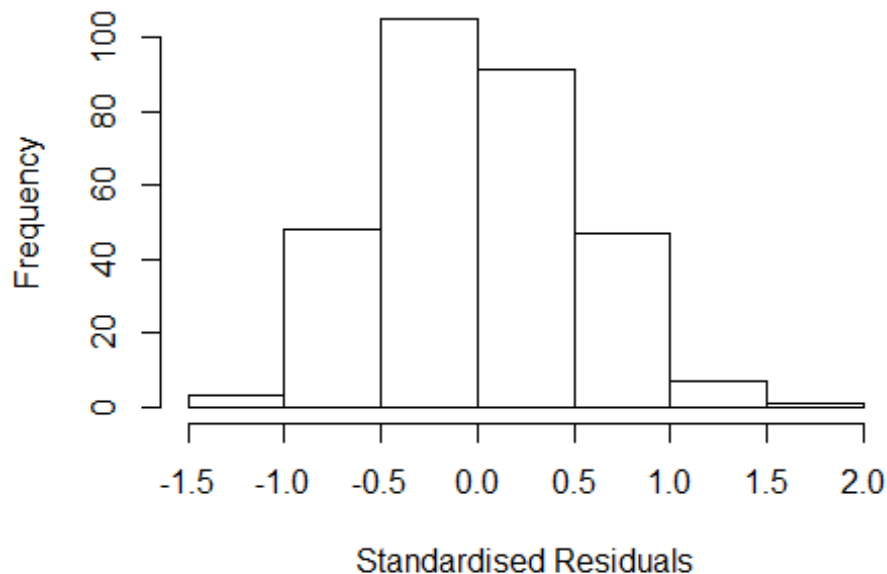
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160, p-value = 0.7
## 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	1.631	1	1.277
LastAuthorFemale	1.722	1	1.312
UniqueAuthors	1.445	3	1.063
Year	1.962	16	1.021

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.2221 -0.3697 -0.0302 0.3780 1.6910
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9679 0.2100 4.61 6.2e-06 ***
## FirstAuthorFemale1 0.0111 0.0968 0.11 0.9089
## LastAuthorFemale1 -0.0152 0.1093 -0.14 0.8895
## UniqueAuthors2 0.2693 0.0714 3.77 0.0002 ***
## UniqueAuthors3 0.2412 0.0946 2.55 0.0113 *
## UniqueAuthors4 0.1435 0.3429 0.42 0.6759
## Year1997 0.5701 0.2575 2.21 0.0276 *
## Year1998 -0.2268 0.2401 -0.94 0.3457
## Year1999 -0.2917 0.2751 -1.06 0.2899
## Year2000 -0.4099 0.2796 -1.47 0.1438
```

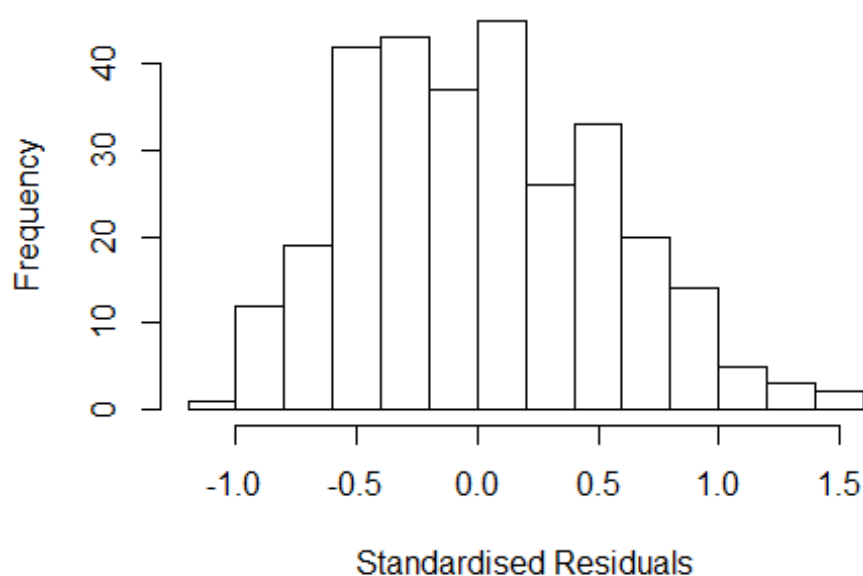


```

## Year2001          -0.1902      0.2503   -0.76   0.4481
## Year2002          -0.1462      0.2912   -0.50   0.6160
## Year2003          -0.2612      0.2434   -1.07   0.2843
## Year2004          -0.3842      0.2374   -1.62   0.1067
## Year2005          -0.2868      0.2437   -1.18   0.2403
## Year2006          -0.2751      0.2314   -1.19   0.2355
## Year2007          -0.3220      0.2323   -1.39   0.1668
## Year2008          -0.1752      0.2349   -0.75   0.4565
## Year2009          -0.2697      0.2396   -1.13   0.2613
## Year2010          -0.2698      0.2328   -1.16   0.2475
## Year2011          -0.3528      0.2670   -1.32   0.1874
## Year2012          -0.4525      0.2515   -1.80   0.0730 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.543
## Multiple R-squared:  0.122, Adjusted R-squared:  0.0566
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 279 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.311  0.890  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      3.31e-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 1.401 1 1.184
## LastAuthorFemale 1.561 1 1.249
## Year 1.440 16 1.011

```

Residuals from first and last author



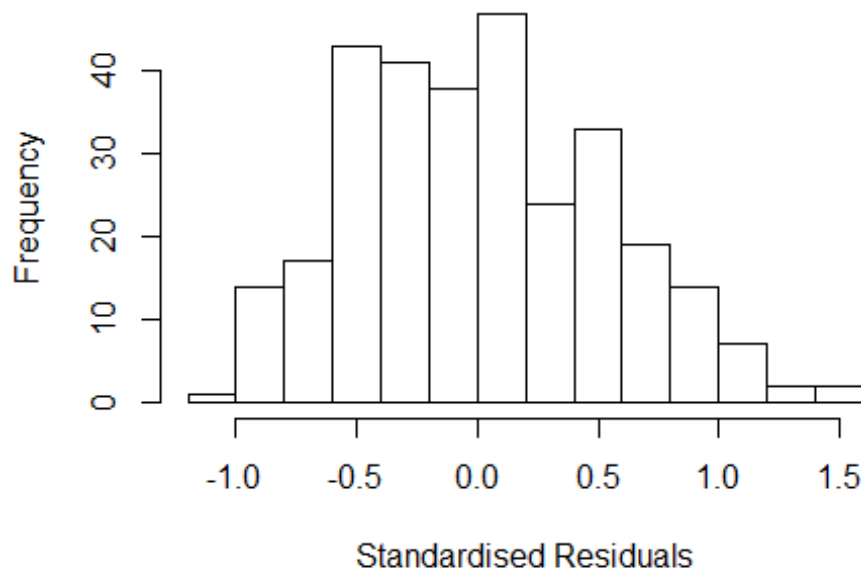
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0540 -0.3950 -0.0121 0.4101 1.5320
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0896 0.2007 5.43 1.2e-07 ***
## FirstAuthorFemale1 0.0472 0.0948 0.50 0.619
## LastAuthorFemale1 -0.0356 0.1103 -0.32 0.747
## Year1997 0.5259 0.2505 2.10 0.037 *
## Year1998 -0.2506 0.2351 -1.07 0.287
## Year1999 -0.3685 0.2660 -1.39 0.167
## Year2000 -0.4134 0.2726 -1.52 0.131
## Year2001 -0.2074 0.2393 -0.87 0.387
## Year2002 -0.1893 0.2890 -0.65 0.513
## Year2003 -0.2340 0.2332 -1.00 0.316
## Year2004 -0.3647 0.2285 -1.60 0.112
## Year2005 -0.2336 0.2354 -0.99 0.322
```

```

## Year2006          -0.2676      0.2210    -1.21     0.227
## Year2007          -0.3202      0.2221    -1.44     0.151
## Year2008          -0.1751      0.2317    -0.76     0.451
## Year2009          -0.2156      0.2316    -0.93     0.353
## Year2010          -0.2398      0.2234    -1.07     0.284
## Year2011          -0.3233      0.2592    -1.25     0.213
## Year2012          -0.3773      0.2369    -1.59     0.112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.0713, Adjusted R-squared:  0.0122
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 275 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.422  0.886   0.945   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      3.31e-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 1.15 1      1.072
## Year              1.15 16      1.004

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.0889 -0.3993 -0.0139 0.4132 1.5349
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0889 0.2017 5.40 1.4e-07 ***
## FirstAuthorFemale1 0.0323 0.0861 0.38 0.707
## Year1997 0.5257 0.2496 2.11 0.036 *
## Year1998 -0.2485 0.2362 -1.05 0.294
## Year1999 -0.3692 0.2658 -1.39 0.166
## Year2000 -0.4155 0.2743 -1.51 0.131
## Year2001 -0.2052 0.2389 -0.86 0.391
## Year2002 -0.1865 0.2896 -0.64 0.520
## Year2003 -0.2380 0.2316 -1.03 0.305
## Year2004 -0.3682 0.2279 -1.62 0.107
## Year2005 -0.2364 0.2370 -1.00 0.319
## Year2006 -0.2698 0.2218 -1.22 0.225
```

```

## Year2007          -0.3257      0.2200   -1.48    0.140
## Year2008          -0.1819      0.2274   -0.80    0.424
## Year2009          -0.2188      0.2314   -0.95    0.345
## Year2010          -0.2417      0.2235   -1.08    0.280
## Year2011          -0.3207      0.2633   -1.22    0.224
## Year2012          -0.3775      0.2391   -1.58    0.115
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.555
## Multiple R-squared:  0.0705, Adjusted R-squared:  0.0149
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 276 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.425  0.886  0.946  0.917  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      3.31e-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 1.269 1      1.127
## Year      1.269 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
## -1.0847 -0.4026 -0.0113  0.4053  1.5382

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0976    0.2007    5.47 9.9e-08 ***
## LastAuthorFemale1 -0.0129    0.0996   -0.13  0.897
## Year1997          0.5294    0.2535    2.09  0.038 *
## Year1998         -0.2546    0.2365   -1.08  0.282
## Year1999         -0.3697    0.2643   -1.40  0.163
## Year2000         -0.4180    0.2724   -1.53  0.126
## Year2001         -0.2011    0.2424   -0.83  0.407
## Year2002         -0.1900    0.2926   -0.65  0.517
## Year2003         -0.2406    0.2343   -1.03  0.305
## Year2004         -0.3676    0.2304   -1.60  0.112
## Year2005         -0.2367    0.2367   -1.00  0.318
## Year2006         -0.2751    0.2214   -1.24  0.215
## Year2007         -0.3177    0.2258   -1.41  0.161
## Year2008         -0.1811    0.2332   -0.78  0.438
## Year2009         -0.2170    0.2332   -0.93  0.353
## Year2010         -0.2447    0.2244   -1.09  0.276
## Year2011         -0.3209    0.2631   -1.22  0.224
## Year2012         -0.3799    0.2388   -1.59  0.113
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.0708, Adjusted R-squared:  0.0152
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 277 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.414  0.883   0.946   0.915   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      3.31e-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: 302"
## [1] ""

```

```

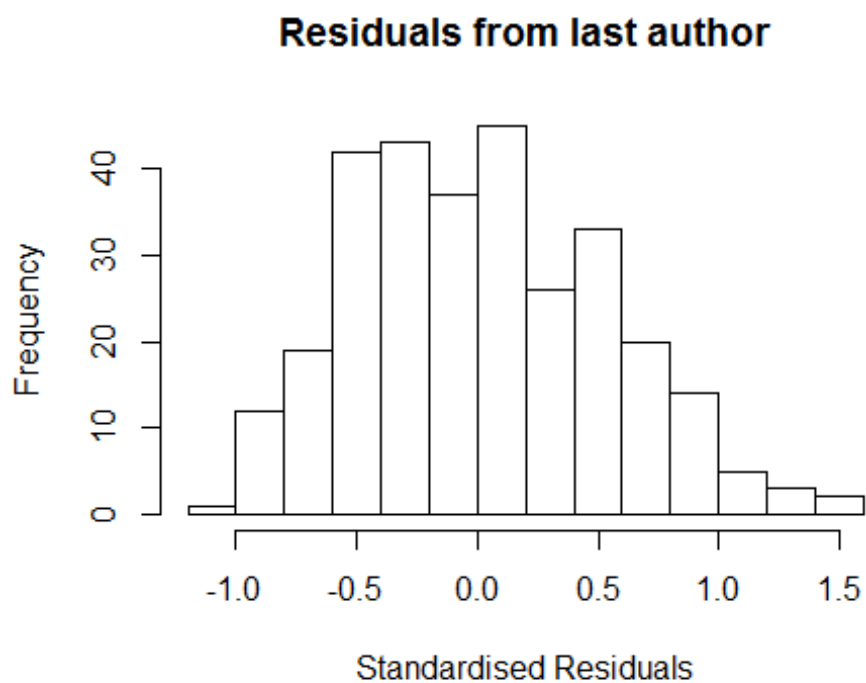
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2100"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    3    2    2    2    1    1    3    2   12    9   18   19   19   11
##
## 1996 1997 1998 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    3    1    1    1    1    1    2    0    5    4   11   15   14    7
##
## 1996 1997 1998 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    3    1    1    1    1    1    2    0    4    4   10   13   11    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.47602664488645"
## [1] "Male first author team size 2018 geometric mean: 2.51984209978975"

## 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 = 6, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 2.9515463231891"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

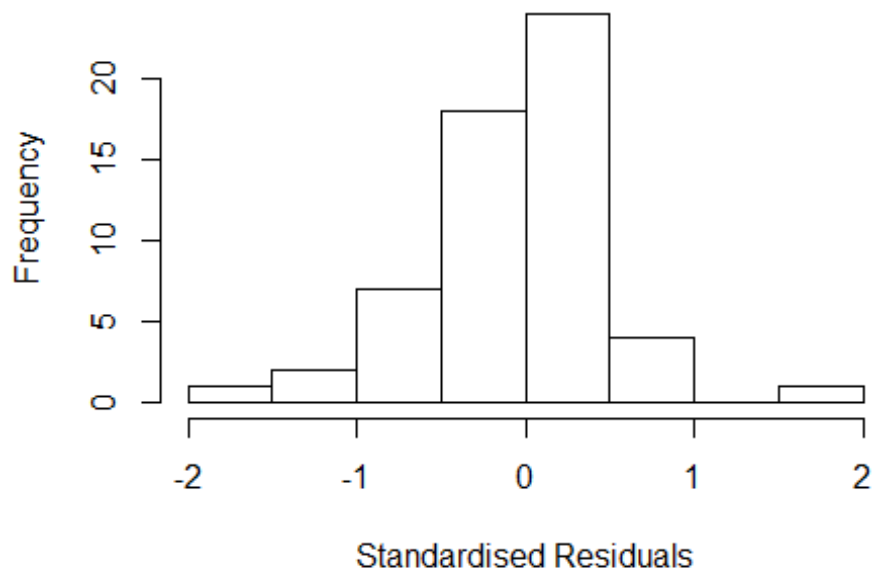
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, 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	1.187e+01	1	3.445
## LastAuthorFemale	7.560e+00	1	2.750
## UniqueAuthors	-1.864e+16	4	NaN
## Year	-3.053e+16	12	NaN

Residuals from first and last author and team size



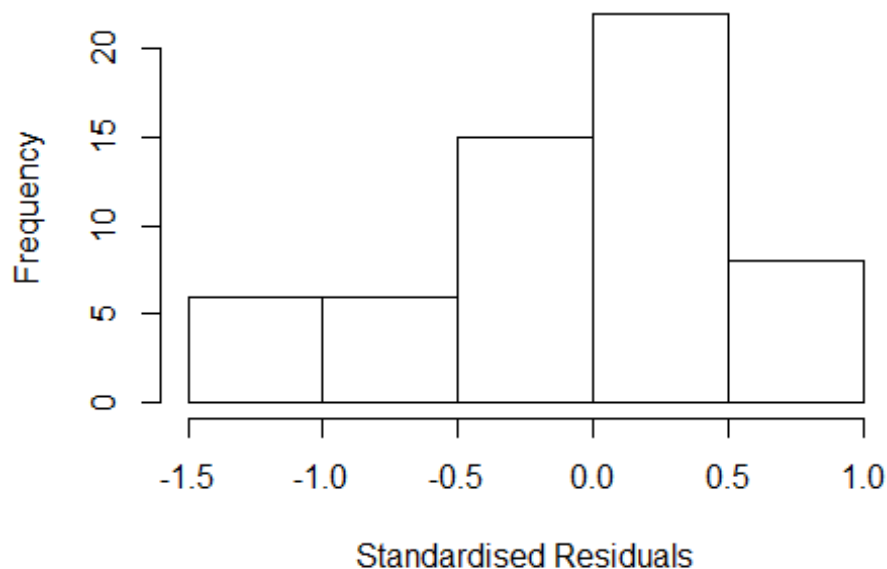
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.8227 -0.1996 0.0273 0.2933 1.5430
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2145 0.7446 1.63 0.1111
## FirstAuthorFemale1 0.0930 0.1994 0.47 0.6436
## LastAuthorFemale1 0.0239 0.1737 0.14 0.8914
## UniqueAuthors2 0.8154 0.4273 1.91 0.0639 .
## UniqueAuthors3 0.6082 0.6286 0.97 0.3394
## UniqueAuthors4 0.7229 0.4983 1.45 0.1550
## UniqueAuthors5 1.1433 0.3626 3.15 0.0032 **
## Year1998 -2.0299 1.1176 -1.82 0.0772 .
## Year2000 -1.2145 0.7446 -1.63 0.1111
## Year2002 0.0445 0.7446 0.06 0.9527
```

```

## Year2003          -0.6065      0.7446   -0.81    0.4204
## Year2004          -1.6879      1.1176   -1.51    0.1392
## Year2005          -1.5003      1.2322   -1.22    0.2309
## Year2007          -0.6816      1.0963   -0.62    0.5378
## Year2008          -0.5382      0.8420   -0.64    0.5265
## Year2009          -0.7365      1.1040   -0.67    0.5087
## Year2010          -0.8164      1.0908   -0.75    0.4588
## Year2011          -0.6469      1.1594   -0.56    0.5801
## Year2012          -0.5629      1.1481   -0.49    0.6268
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.463
## Multiple R-squared:  0.491, Adjusted R-squared:  0.25
## Convergence in 41 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 49 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0873 0.8830 0.9510 0.8910 0.9900 0.9990
## Algorithmic parameters:
##           tuning.chi          bb          tuning.psi          refine.tol
##           1.55e+00          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-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 1.678 1          1.296
## LastAuthorFemale  4.015 1          2.004
## Year              8.768 12          1.095

```

Residuals from first and last author



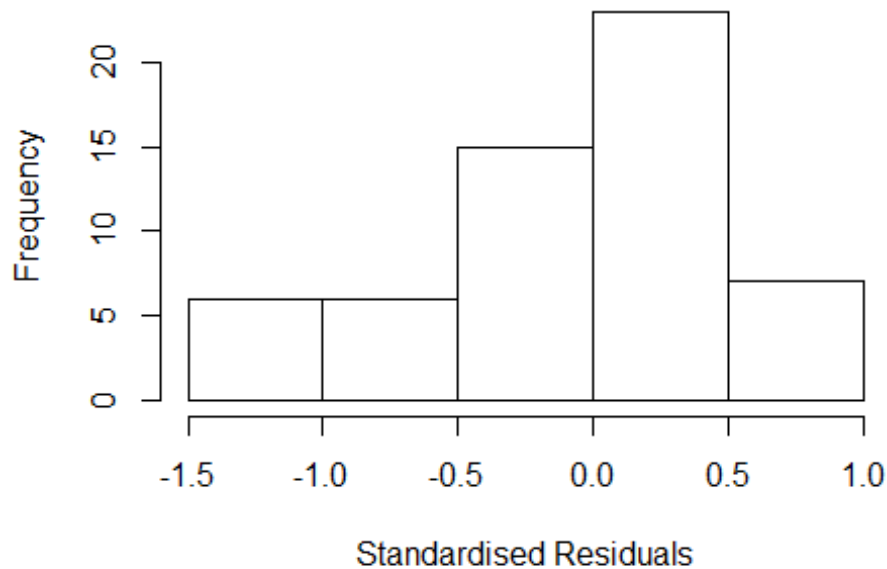
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2092 -0.1109 0.0517 0.2836 0.8224
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.91361 0.47213 1.94 0.06 .
## FirstAuthorFemale1 0.07661 0.23357 0.33 0.74
## LastAuthorFemale1 -0.00575 0.23054 -0.02 0.98
## Year1998 -0.91361 0.47213 -1.94 0.06 .
## Year2000 -0.91361 0.47213 -1.94 0.06 .
## Year2002 0.34539 0.47213 0.73 0.47
## Year2003 -0.30561 0.47213 -0.65 0.52
## Year2004 -0.57161 0.47213 -1.21 0.23
## Year2005 -0.48761 0.47704 -1.02 0.31
## Year2007 0.03850 0.59477 0.06 0.95
## Year2008 0.09488 0.58137 0.16 0.87
## Year2009 0.28496 0.52270 0.55 0.59
```

```

## Year2010          0.19073    0.51919    0.37    0.72
## Year2011          0.26744    0.51775    0.52    0.61
## Year2012          0.29564    0.62061    0.48    0.64
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.547
## Multiple R-squared:  0.225, Adjusted R-squared:  -0.034
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.604  0.822  0.957  0.893  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      1.75e-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 6.432 1          2.536
## Year              6.432 12          1.081

```

Residuals from first author



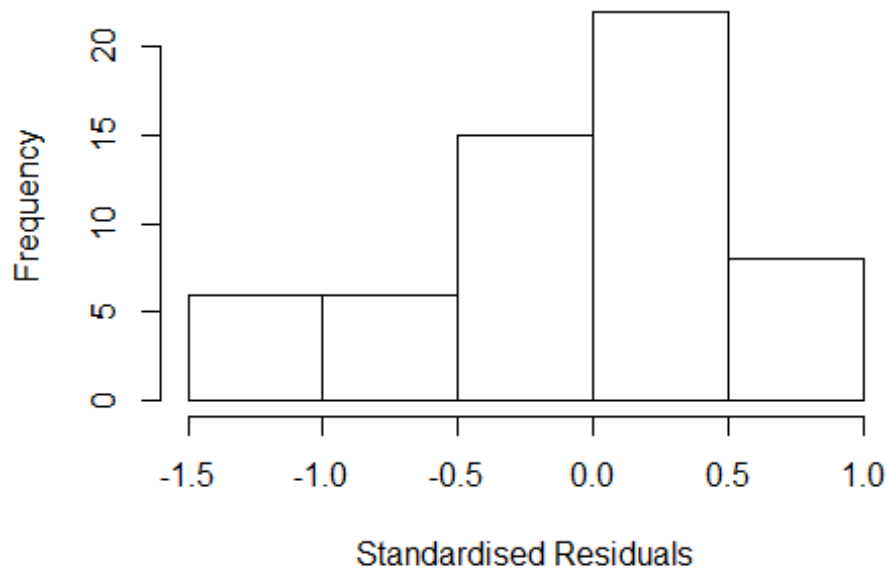
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.213 -0.118 0.045 0.278 0.823
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9185 0.4814 1.91 0.063 .
## FirstAuthorFemale1 0.0763 0.2491 0.31 0.761
## Year1998 -0.9185 0.4814 -1.91 0.063 .
## Year2000 -0.9185 0.4814 -1.91 0.063 .
## Year2002 0.3405 0.4814 0.71 0.483
## Year2003 -0.3105 0.4814 -0.65 0.522
## Year2004 -0.5765 0.4814 -1.20 0.238
## Year2005 -0.4925 0.4862 -1.01 0.317
## Year2007 0.0339 0.5800 0.06 0.954
## Year2008 0.0965 0.5925 0.16 0.871
## Year2009 0.2798 0.5259 0.53 0.597
## Year2010 0.1930 0.5283 0.37 0.717
```

```

## Year2011          0.2635      0.4987      0.53      0.600
## Year2012          0.2947      0.5863      0.50      0.618
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.528
## Multiple R-squared:  0.23,   Adjusted R-squared:  -0.00282
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.577  0.810   0.955   0.885   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      1.75e-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 4.475 1          2.115
## Year             4.475 12          1.064

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2497 -0.0960 0.0264 0.2924 0.8224
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9187 0.4817 1.91 0.063 .
## LastAuthorFemale1 0.0032 0.2412 0.01 0.989
## Year1998 -0.9187 0.4817 -1.91 0.063 .
## Year2000 -0.9187 0.4817 -1.91 0.063 .
## Year2002 0.3403 0.4817 0.71 0.484
## Year2003 -0.3107 0.4817 -0.65 0.522
## Year2004 -0.5767 0.4817 -1.20 0.238
## Year2005 -0.4927 0.4865 -1.01 0.317
## Year2007 0.0321 0.6043 0.05 0.958
## Year2008 0.0985 0.5828 0.17 0.867
## Year2009 0.3146 0.5093 0.62 0.540
## Year2010 0.2059 0.5272 0.39 0.698
```

```

## Year2011          0.2788      0.5197      0.54      0.594
## Year2012          0.3311      0.6142      0.54      0.593
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.527
## Multiple R-squared:  0.229, Adjusted R-squared:  -0.004
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 45 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.554  0.793  0.955  0.882  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          1.75e-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:  57"
## [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]"
##
## 2005 2006 2012
##    2    1    1
##
## 2005 2006 2012
##    1    1    0
##
## 2005 2006 2012
##    1    1    0
## [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: 2"

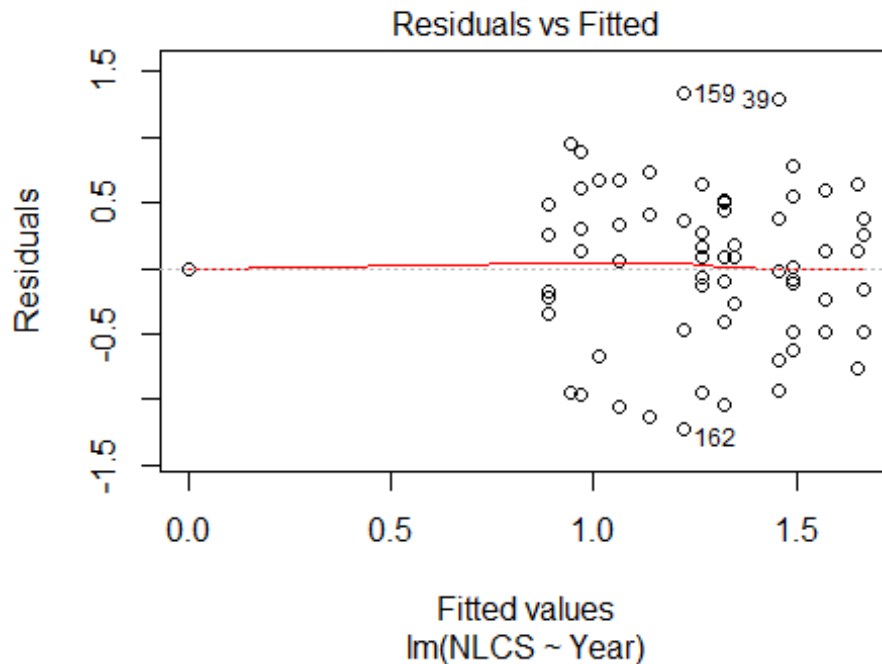
```



```

## [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 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
## 14 10 15 1 18 13 8 11 12 10 5 10 11 7 6
## 2011 2012
## 10 8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 2 5 0 7 2 4 7 7 4 2 6 5 3 3
## 2011 2012
## 4 4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 2 4 0 7 2 3 3 6 3 2 5 4 3 3
## 2011 2012
## 3 4
## [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 = 15, p-value <2e-16

```



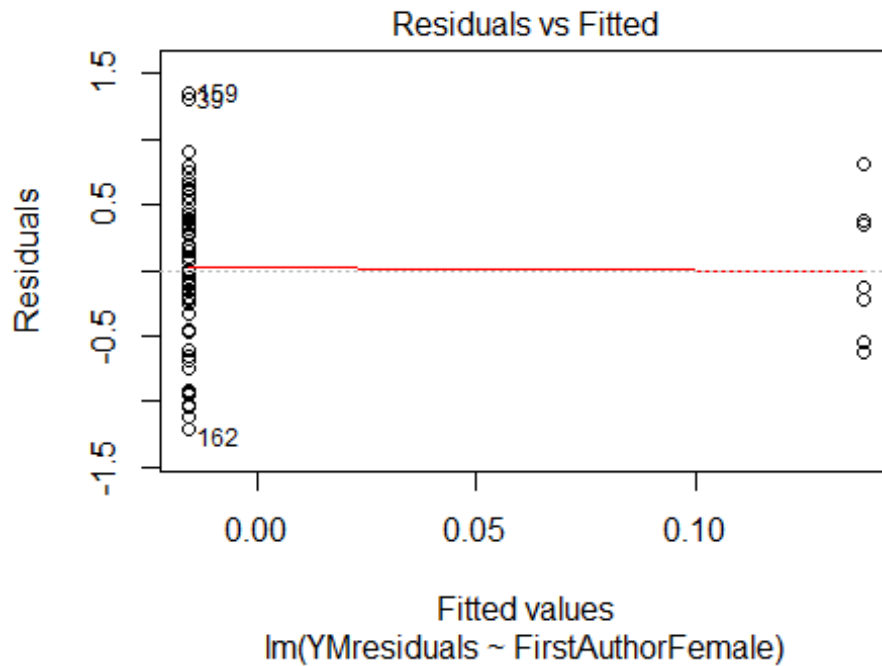
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.2, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male first author team size 2018 geometric mean: 2.24492409661875"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

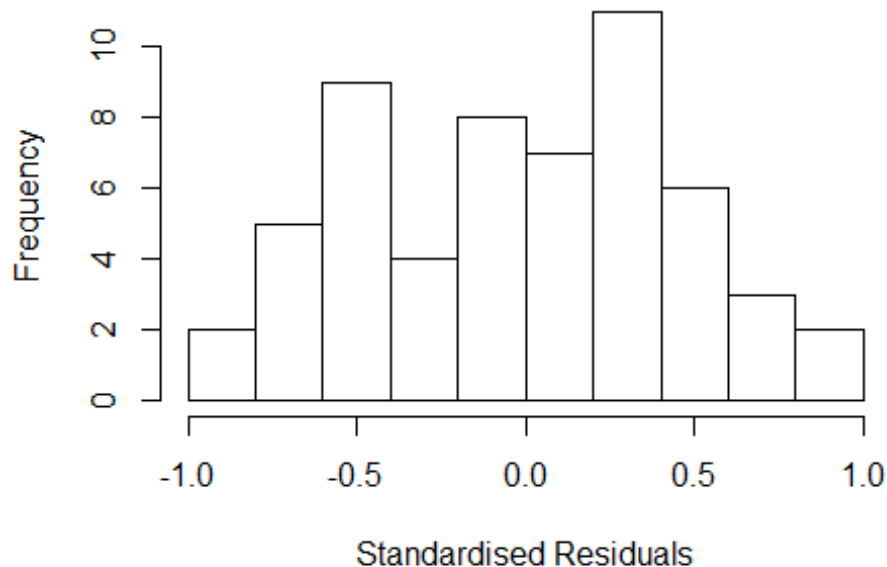
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 2.20817902734762"

## 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, 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 3.196e+00 1      1.788
## LastAuthorFemale  4.926e+00 1      2.219
## UniqueAuthors    8.502e+15 4     97.991
## Year             4.315e+16 15     3.585
```

Residuals from first and last author and team size



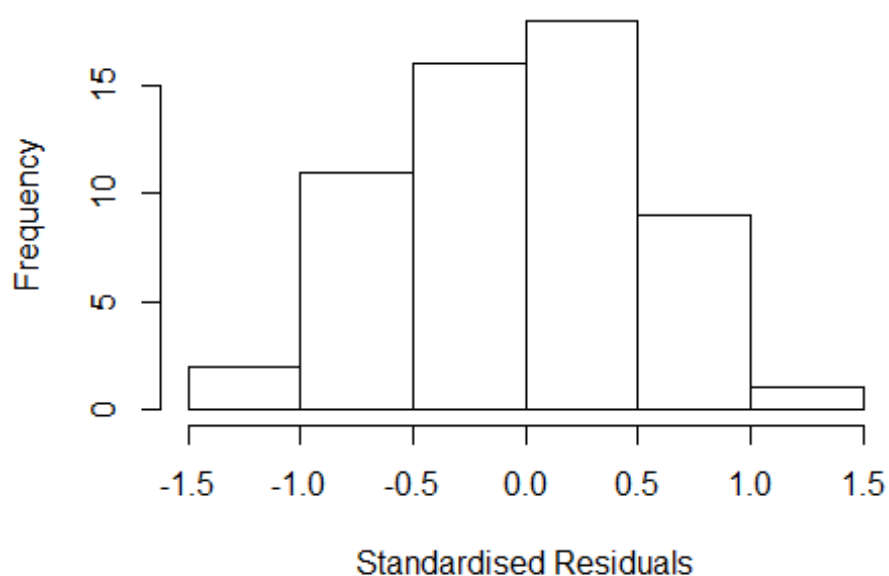
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8811 -0.4139 0.0372 0.3531 0.9589
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.63411 0.37862 1.67 0.1029
## FirstAuthorFemale1 1.08465 0.35729 3.04 0.0045 **
## LastAuthorFemale1 -0.20708 0.41356 -0.50 0.6197
## UniqueAuthors2 0.77531 0.23275 3.33 0.0020 **
## UniqueAuthors3 0.00328 0.28434 0.01 0.9909
## UniqueAuthors4 -0.15559 0.46076 -0.34 0.7376
## UniqueAuthors5 -0.38077 0.23487 -1.62 0.1140
## Year1997 -0.63411 0.37862 -1.67 0.1029
## Year1998 0.64035 0.53908 1.19 0.2429
## Year2000 0.42860 0.38194 1.12 0.2694
```

```

## Year2001      -0.15628    0.49008   -0.32    0.7517
## Year2002      -0.09933    0.38091   -0.26    0.7958
## Year2003       0.82716    0.45944    1.80    0.0804 .
## Year2004      -0.04265    0.40996   -0.10    0.9177
## Year2005      -0.05921    0.37106   -0.16    0.8741
## Year2006       0.37611    0.76809    0.49    0.6274
## Year2007      -0.16042    0.42092   -0.38    0.7054
## Year2008      -0.31519    0.34301   -0.92    0.3644
## Year2009       1.13076    0.63002    1.79    0.0813 .
## Year2010       0.52650    0.40317    1.31    0.2001
## Year2011       1.14427    0.42312    2.70    0.0105 *
## Year2012       0.92966    0.40961    2.27    0.0295 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.522, Adjusted R-squared:  0.236
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 50 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.742  0.904  0.949  0.929  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      1.75e-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 2.841 1 1.685
## LastAuthorFemale 11.543 1 3.398
## Year 22.092 15 1.109

```

Residuals from first and last author



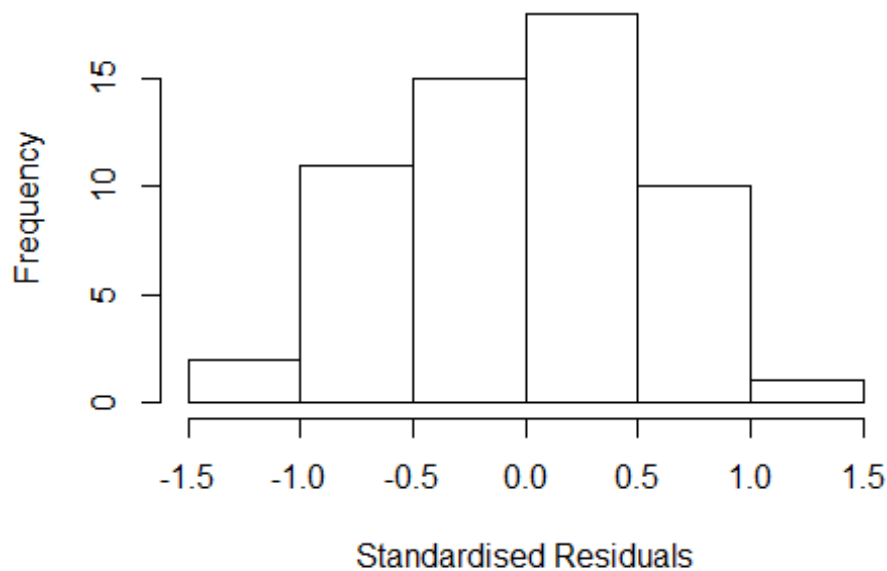
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.25e+00 -4.58e-01 -2.22e-16 3.77e-01 1.36e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2493 0.6111 2.04 0.048 *
## FirstAuthorFemale1 0.6893 0.3780 1.82 0.076 .
## LastAuthorFemale1 -0.2927 0.3134 -0.93 0.356
## Year1997 -1.2493 0.6111 -2.04 0.048 *
## Year1998 0.1317 0.9209 0.14 0.887
## Year2000 0.0394 0.6281 0.06 0.950
## Year2001 -0.6499 0.8067 -0.81 0.425
## Year2002 -0.1255 0.7915 -0.16 0.875
## Year2003 0.2170 0.6728 0.32 0.749
## Year2004 -0.1242 0.6481 -0.19 0.849
## Year2005 -0.4774 0.7605 -0.63 0.534
## Year2006 -0.2358 0.8699 -0.27 0.788
```

```

## Year2007          -0.4368      0.6949   -0.63    0.533
## Year2008          -0.4458      0.6322   -0.71    0.485
## Year2009           0.5382      0.7572    0.71    0.481
## Year2010           0.1946      0.6170    0.32    0.754
## Year2011           0.4811      0.6454    0.75    0.460
## Year2012           0.4186      0.6340    0.66    0.513
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.641
## Multiple R-squared:  0.32,   Adjusted R-squared:  0.0239
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.630  0.903   0.953   0.926   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-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 2.265 1          1.505
## Year              2.265 15          1.028

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2426 -0.4585 0.0652 0.3773 1.3587
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2426 0.6052 2.05 0.047 *
## FirstAuthorFemale1 0.6310 0.3199 1.97 0.055 .
## Year1997 -1.2426 0.6052 -2.05 0.047 *
## Year1998 0.1436 0.9059 0.16 0.875
## Year2000 0.0452 0.6226 0.07 0.943
## Year2001 -0.6142 0.8134 -0.76 0.455
## Year2002 -0.1233 0.7854 -0.16 0.876
## Year2003 0.2239 0.6663 0.34 0.739
## Year2004 -0.2013 0.6542 -0.31 0.760
## Year2005 -0.4705 0.7534 -0.62 0.536
## Year2006 -0.2291 0.8600 -0.27 0.791
## Year2007 -0.4313 0.6891 -0.63 0.535
```



```

## Year2008          -0.4242      0.6205   -0.68    0.498
## Year2009          0.4218      0.7017    0.60    0.551
## Year2010          0.1046      0.6146    0.17    0.866
## Year2011          0.4878      0.6396    0.76    0.450
## Year2012          0.4251      0.6285    0.68    0.503
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.308, Adjusted R-squared:  0.0319
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.646  0.888  0.954  0.928  0.987  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      1.75e-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 6.048 1          2.459
## Year              6.048 15          1.062

## [1] "List of 0 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
## -1.2478 -0.3984  0.0267  0.4316  1.3628
##

```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2478    0.6077   2.05  0.047 *
## LastAuthorFemale1 -0.1855    0.3848  -0.48  0.632
## Year1997         -1.2478    0.6077  -2.05  0.047 *
## Year1998          0.1344    0.9119   0.15  0.884
## Year2000          0.0407    0.6254   0.07  0.948
## Year2001         -0.3038    1.3103  -0.23  0.818
## Year2002         -0.1250    0.7901  -0.16  0.875
## Year2003          0.2186    0.6688   0.33  0.746
## Year2004          0.0571    0.6354   0.09  0.929
## Year2005         -0.4758    0.7565  -0.63  0.533
## Year2006         -0.2343    0.8388  -0.28  0.781
## Year2007         -0.4356    0.6932  -0.63  0.533
## Year2008         -0.2713    0.6279  -0.43  0.668
## Year2009          0.4948    0.7576   0.65  0.517
## Year2010          0.1605    0.6190   0.26  0.797
## Year2011          0.4827    0.6421   0.75  0.457
## Year2012          0.4201    0.6311   0.67  0.509
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.277, Adjusted R-squared:  -0.0125
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~ = 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.634  0.874  0.954  0.922  0.985  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      1.75e-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: 57"
## [1] ""
## [1] ""
## [1] "#####"

```

```

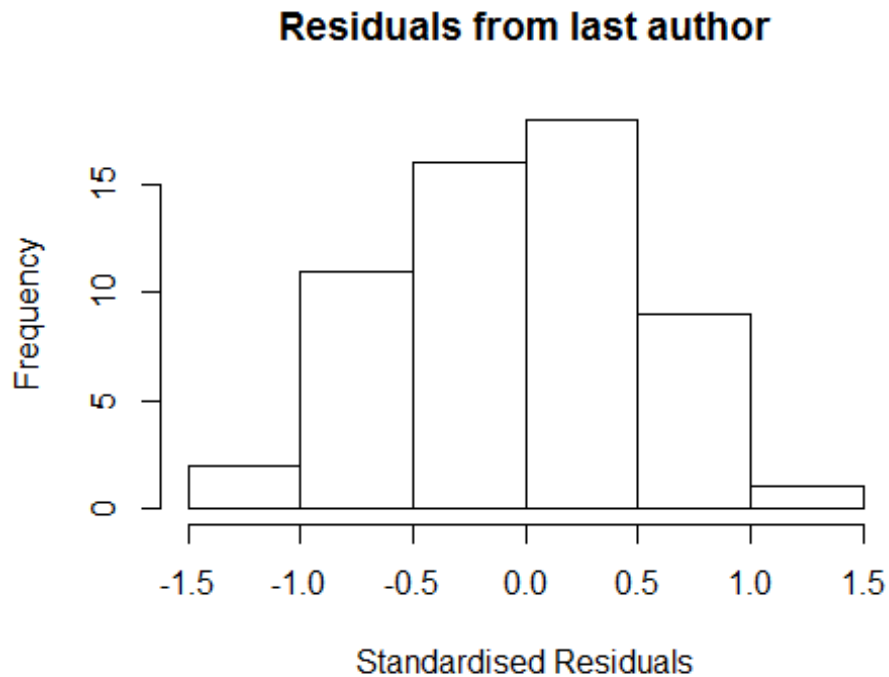
## [1] "Analysis of AJSC 2103"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      8    4   11    6    5   10    2    5    6    2    3    1    5    5    3
## 2011 2012
##      4    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      3    0    3    2    3    2    2    3    3    0    1    1    3    4    0
## 2011 2012
##      2    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      3    0    3    2    3    2    1    3    3    0    1    1    3    4    0
## 2011 2012
##      1    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male first author team size 2018 geometric mean: 2.51984209978975"

## 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 = 1, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 2.37841423000544"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

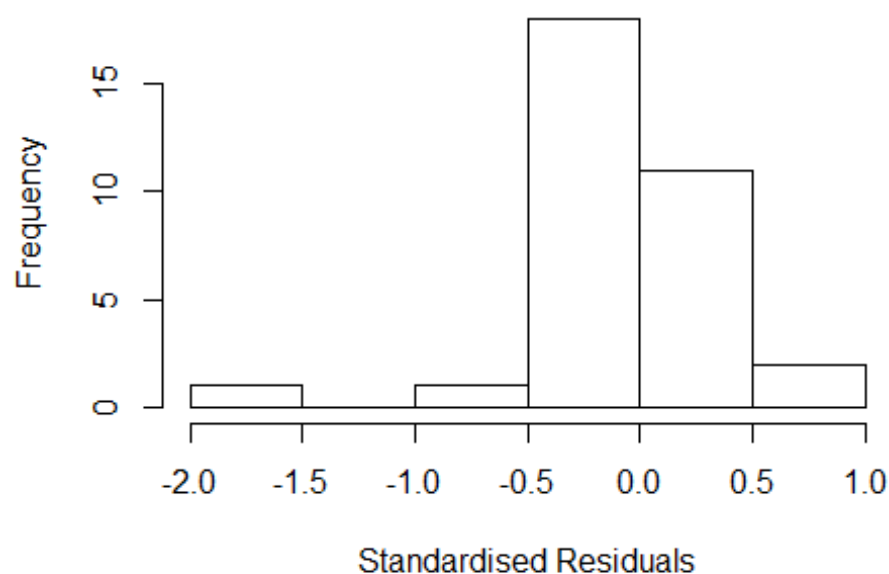
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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"
## [1] "Regression 2: First author gender, Last author gender, Year as
##      factors"
##
```

	GVIF	Df	GVIF ^{1/(2*Df)}
FirstAuthorFemale	1.178e+13	1	3432773
LastAuthorFemale	-7.586e+13	1	NaN
Year	-7.520e+26	13	NaN

Residuals from first and last author



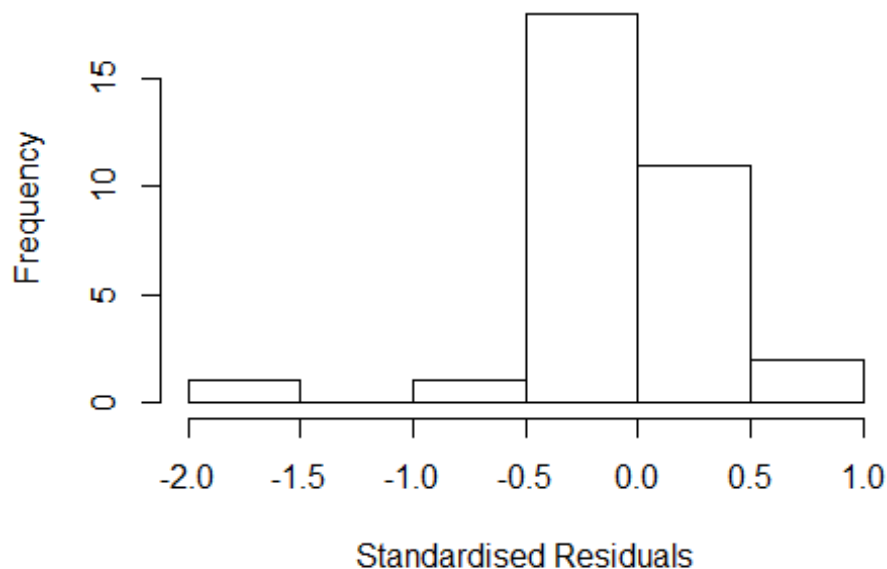
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.60e+00 -1.67e-01 -6.66e-16  1.43e-01  6.84e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.59953    0.23099   6.92 2.5e-06 ***
## FirstAuthorFemale1 -0.00859    0.12320  -0.07  0.9452
## LastAuthorFemale1 -0.28361    0.09395  -3.02  0.0077 **
## Year1998        -0.44697    0.50006  -0.89  0.3839
## Year1999        -0.29623    0.45060  -0.66  0.5197
## Year2000        -0.17139    0.32206  -0.53  0.6015
## Year2001        -0.09273    0.30124  -0.31  0.7620
## Year2002         0.13647    0.23099   0.59  0.5624
## Year2003        -0.14363    0.45400  -0.32  0.7556
## Year2004        -0.13977    0.29764  -0.47  0.6446
## Year2006        -0.70153    0.23099  -3.04  0.0074 **
## Year2007       -1.59953    0.23099  -6.92 2.5e-06 ***
```

```

## Year2008          -0.12759    0.33357   -0.38    0.7068
## Year2009          -0.01732    0.26109   -0.07    0.9479
## Year2011           0.56247    0.23099    2.44    0.0262 *
## Year2012          -0.23144    0.27512   -0.84    0.4119
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.425
## Multiple R-squared:  0.519, Adjusted R-squared:  0.0951
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.126  0.936  0.981  0.921  0.993  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      3.03e-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 5.578e-12 1      2.362e-06
## Year              5.578e-12 13      3.691e-01

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.58e+00 -1.85e-01 -1.05e-15 1.62e-01 6.83e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5823 0.3216 4.92 0.00011 ***
## FirstAuthorFemale1 -0.0636 0.1192 -0.53 0.60020
## Year1998 -0.4284 0.5422 -0.79 0.43968
## Year1999 -0.2515 0.3260 -0.77 0.45039
## Year2000 -0.1553 0.3939 -0.39 0.69796
## Year2001 -0.0480 0.3821 -0.13 0.90139
## Year2002 0.1537 0.3216 0.48 0.63851
## Year2003 -0.1259 0.5007 -0.25 0.80435
## Year2004 -0.1225 0.3697 -0.33 0.74421
## Year2006 -0.6843 0.3216 -2.13 0.04745 *
## Year2007 -1.5823 0.3216 -4.92 0.00011 ***
## Year2008 -0.0553 0.3956 -0.14 0.89029
```

```

## Year2009          -0.0436      0.3418   -0.13  0.89990
## Year2011           0.5797      0.3216    1.80  0.08827 .
## Year2012          -0.1954      0.3478   -0.56  0.58107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.503, Adjusted R-squared:  0.117
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.148  0.948  0.976  0.925  0.993  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      3.03e-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 4.363e+14 1      2.089e+07
## Year              4.363e+14 13      3.657e+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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q    Median      3Q      Max
## -1.64e+00 -1.63e-01 -5.55e-16  1.46e-01  6.88e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)

```



```

## (Intercept)          1.6395      0.1727      9.49 2.0e-08 ***
## LastAuthorFemale1    -0.2893      0.0562     -5.15 6.7e-05 ***
## Year1998              -0.4906      0.4551     -1.08 0.29533
## Year1999              -0.3405      0.3848     -0.88 0.38786
## Year2000              -0.2081      0.3228     -0.64 0.52720
## Year2001              -0.1370      0.2230     -0.61 0.54664
## Year2002               0.0965      0.1727      0.56 0.58308
## Year2003              -0.1850      0.4081     -0.45 0.65569
## Year2004              -0.1799      0.2493     -0.72 0.47987
## Year2006              -0.7415      0.1727     -4.29 0.00044 ***
## Year2007              -1.6395      0.1727     -9.49 2.0e-08 ***
## Year2008              -0.1762      0.2125     -0.83 0.41780
## Year2009              -0.0601      0.1775     -0.34 0.73868
## Year2011               0.5225      0.1727      3.03 0.00726 **
## Year2012              -0.2743      0.1840     -1.49 0.15329
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.537, Adjusted R-squared:  0.178
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0763 0.9320 0.9800 0.9160 0.9920 0.9980
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+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-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: 33"
## [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 2003 2004 2006 2008 2009 2010 2011
##    4    3    3    1    3    2    1    3    1    1    2    3    6
##
## 1996 1997 1998 1999 2000 2001 2003 2004 2006 2008 2009 2010 2011
##    1    0    1    0    2    1    0    2    0    0    0    0    3
##
## 1996 1997 1998 1999 2000 2001 2003 2004 2006 2008 2009 2010 2011
##    1    0    1    0    2    1    0    2    0    0    0    0    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male first author team size 2018 geometric mean: 2"

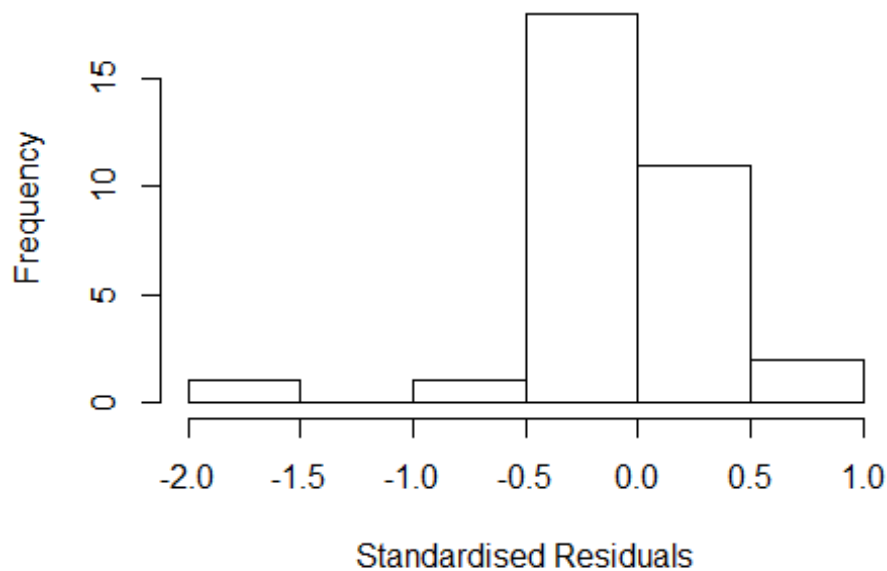
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [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

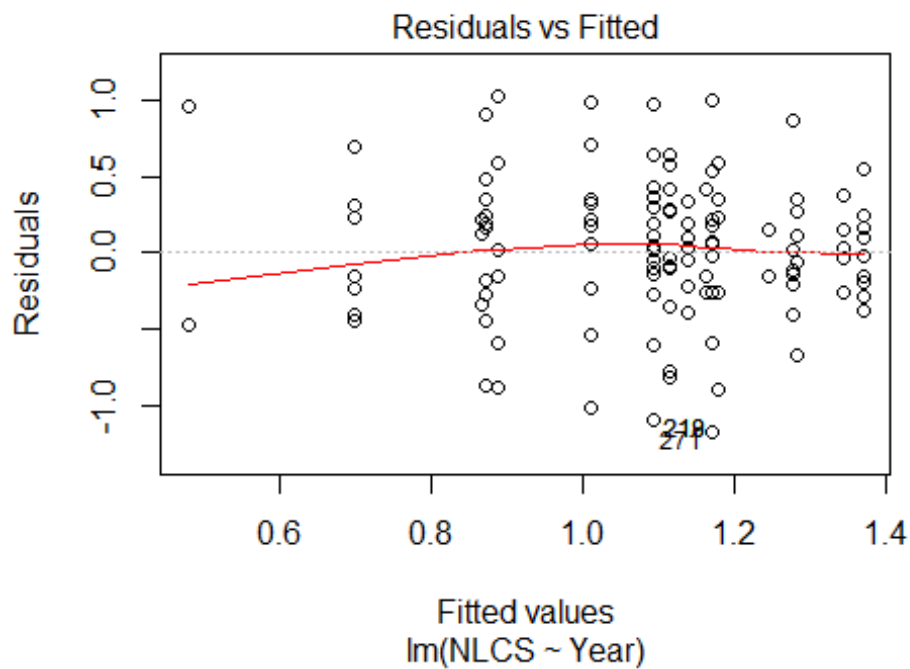
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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,
## [1] "Regression 2: First author gender, Last author gender, Year as
## [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: 9"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2105"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
## [1] "gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 21 11 5 14 17 14 3 10 9 8 11 18 10 30 21
## 2011 2012
## 21 21
##
```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6      3      3      6      9      6      2      3      5      7      7     12      5     20     11
## 2011 2012
##     10     11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6      2      3      6      8      5      2      3      5      7      4     10      5     18     11
## 2011 2012
##      8      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 = 18, df = 16, p-value = 0.3
```

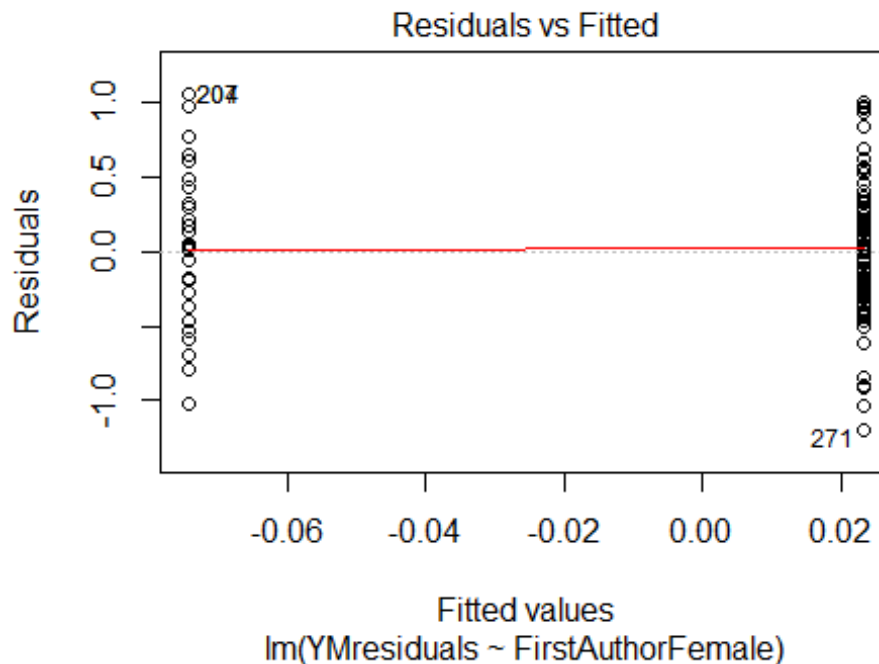


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.5, df = 1, p-value = 0.03
## [1] "Female first author team size 2018 geometric mean: 2.71682189631809"
## [1] "Male first author team size 2018 geometric mean: 2.09306617681083"
```

```
## 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 = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.36297040916991"
## [1] "Male last author team size 2018 geometric mean: 2.24227060938863"

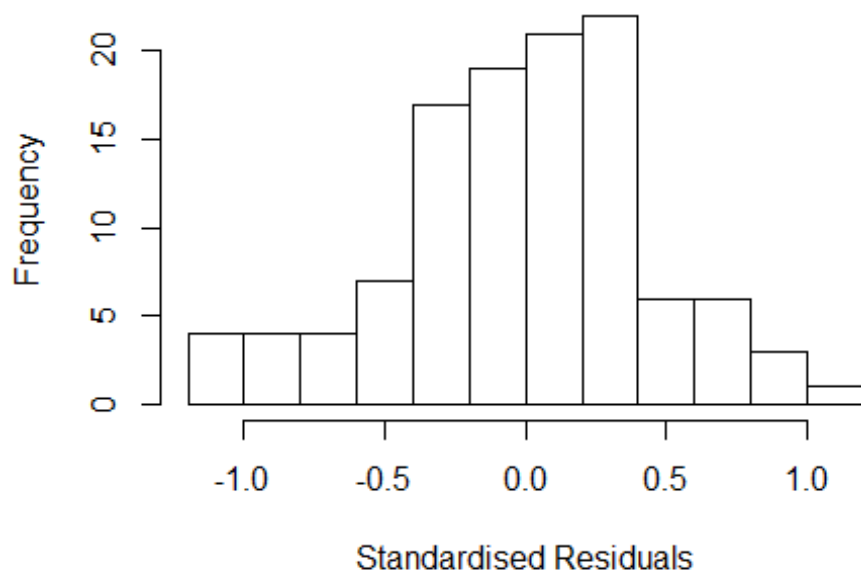
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 94, 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 3.698 1          1.923
```

## LastAuthorFemale	3.660	1	1.913
## UniqueAuthors	35.712	4	1.564
## Year	111.344	16	1.159

Residuals from first and last author and team size



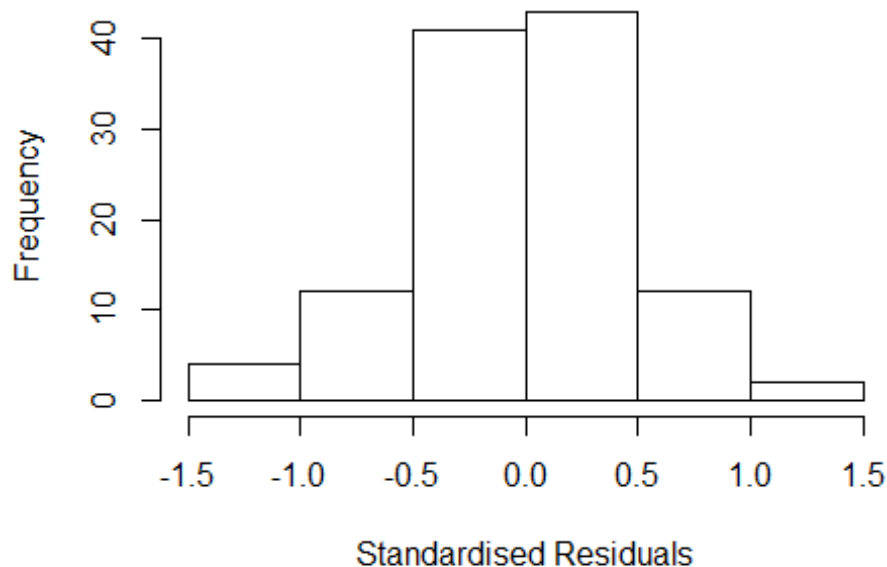
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0883 -0.2859 0.0461 0.2819 1.0462
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.86276 0.39713 2.17 0.032 *
## FirstAuthorFemale1 0.14073 0.15152 0.93 0.355
## LastAuthorFemale1 -0.41833 0.17636 -2.37 0.020 *
## UniqueAuthors2 -0.04902 0.12403 -0.40 0.694
## UniqueAuthors3 0.22553 0.14888 1.51 0.133
## UniqueAuthors4 -0.00853 0.19654 -0.04 0.965
```

```

## UniqueAuthors5      0.37614      0.37963      0.99      0.324
## Year1997             -0.86276      0.39713     -2.17      0.032 *
## Year1998              0.04141      0.42235      0.10      0.922
## Year1999              0.63357      0.42734      1.48      0.142
## Year2000              0.62734      0.43225      1.45      0.150
## Year2001              0.49418      0.44585      1.11      0.271
## Year2002              0.22018      0.52123      0.42      0.674
## Year2003              0.24434      0.46764      0.52      0.603
## Year2004              0.31652      0.47552      0.67      0.507
## Year2005              0.32643      0.42032      0.78      0.439
## Year2006             -0.20050      0.43818     -0.46      0.648
## Year2007              0.14885      0.43813      0.34      0.735
## Year2008              0.31552      0.44686      0.71      0.482
## Year2009              0.24892      0.41220      0.60      0.547
## Year2010              0.17410      0.46788      0.37      0.711
## Year2011              0.24485      0.45855      0.53      0.595
## Year2012              0.32226      0.42137      0.76      0.446
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.473
## Multiple R-squared:  0.266, Adjusted R-squared:  0.0884
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 103 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.575  0.881  0.955  0.914  0.986  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      8.77e-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 3.565 1      1.888
## LastAuthorFemale  3.528 1      1.878
## Year               4.620 16      1.049

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.1843 -0.3162 0.0202 0.2666 1.2154
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8504 0.3469 2.45 0.016 *
## FirstAuthorFemale1 0.1385 0.1577 0.88 0.382
## LastAuthorFemale1 -0.4270 0.1782 -2.40 0.018 *
## Year1997 -0.8504 0.3469 -2.45 0.016 *
## Year1998 0.0197 0.3766 0.05 0.958
## Year1999 0.6407 0.3728 1.72 0.089 .
## Year2000 0.6858 0.3740 1.83 0.070 .
## Year2001 0.5388 0.3891 1.38 0.169
## Year2002 0.3961 0.3634 1.09 0.279
## Year2003 0.3057 0.3912 0.78 0.437
## Year2004 0.3944 0.4093 0.96 0.338
## Year2005 0.3750 0.3691 1.02 0.312
```

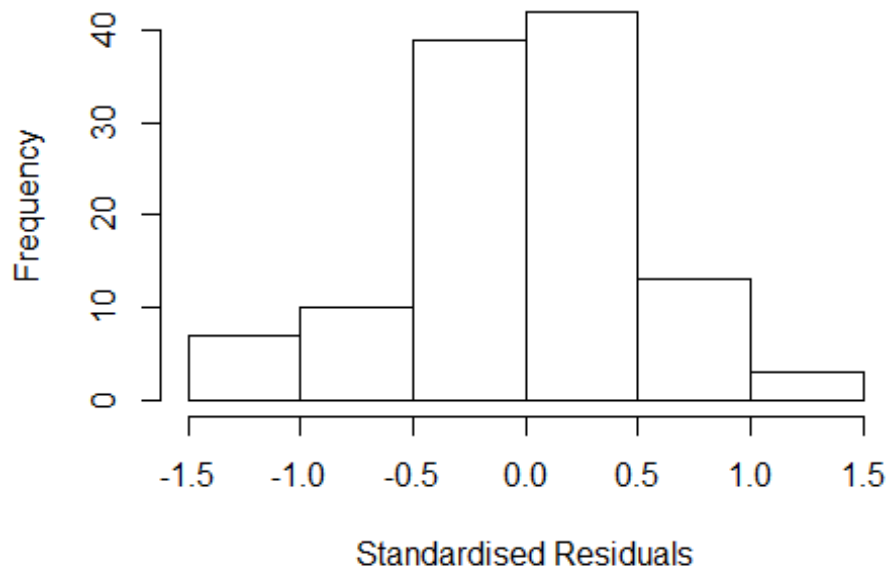


```

## Year2006          -0.0747      0.4056   -0.18    0.854
## Year2007           0.2168      0.3806    0.57    0.570
## Year2008           0.3765      0.3994    0.94    0.348
## Year2009           0.2938      0.3685    0.80    0.427
## Year2010           0.2586      0.4115    0.63    0.531
## Year2011           0.3339      0.4210    0.79    0.430
## Year2012           0.3366      0.3865    0.87    0.386
## ---
## 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:  0.0808
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 102 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.493  0.878  0.959  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      8.77e-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 4.323 1      2.079
## Year              4.323 16      1.047

```

Residuals from first author



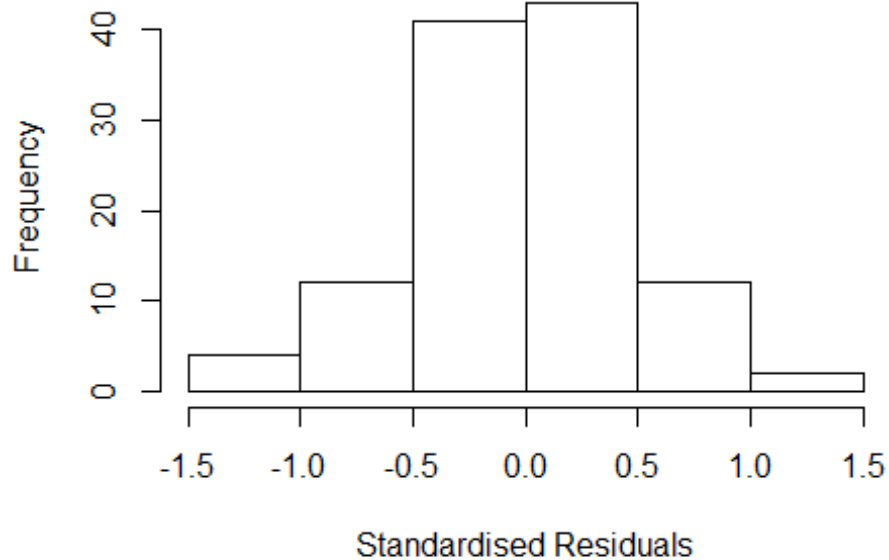
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2144 -0.2714 0.0114 0.2594 1.0573
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8517 0.3407 2.50 0.014 *
## FirstAuthorFemale1 -0.0696 0.1776 -0.39 0.696
## Year1997 -0.8517 0.3407 -2.50 0.014 *
## Year1998 0.0182 0.3706 0.05 0.961
## Year1999 0.5272 0.3650 1.44 0.152
## Year2000 0.5444 0.3578 1.52 0.131
## Year2001 0.4864 0.3845 1.26 0.209
## Year2002 0.3948 0.3575 1.10 0.272
## Year2003 0.3045 0.3861 0.79 0.432
## Year2004 0.3942 0.4281 0.92 0.360
## Year2005 0.3099 0.3570 0.87 0.388
## Year2006 -0.0345 0.4221 -0.08 0.935
```

```

## Year2007          0.1822      0.3694      0.49      0.623
## Year2008          0.4890      0.3837      1.27      0.206
## Year2009          0.2808      0.3728      0.75      0.453
## Year2010          0.2300      0.4109      0.56      0.577
## Year2011          0.3627      0.4229      0.86      0.393
## Year2012          0.3147      0.3813      0.83      0.411
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.482
## Multiple R-squared:  0.174, Adjusted R-squared:  0.0272
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 98 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.505  0.867   0.963   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      8.77e-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 2.577 1      1.605
## Year            2.577 16      1.030

```

Residuals from last author



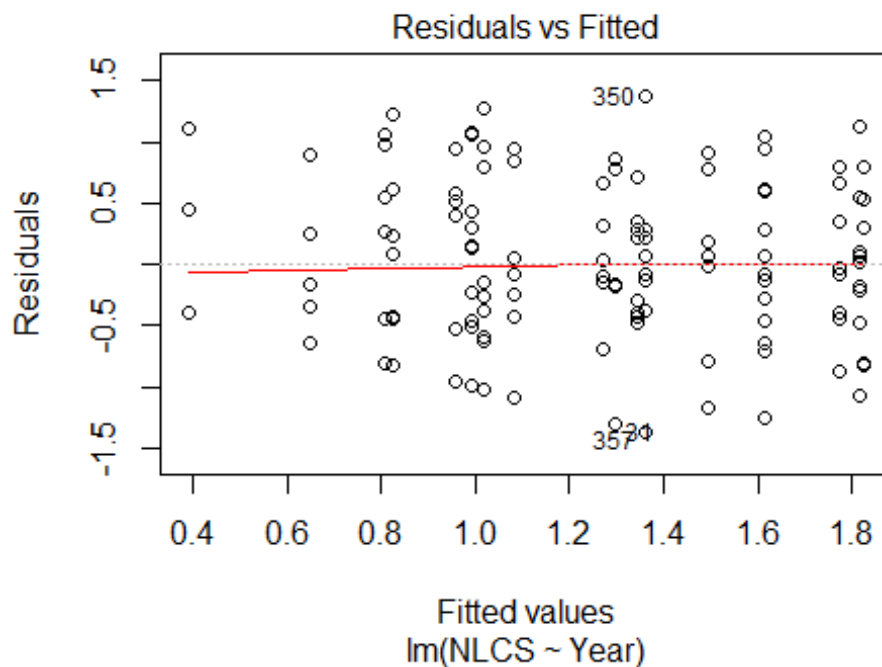
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2055 -0.3220 0.0328 0.2595 1.2611
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8509 0.3457 2.46 0.016 *
## LastAuthorFemale1 -0.3426 0.1532 -2.24 0.028 *
## Year1997 -0.8509 0.3457 -2.46 0.016 *
## Year1998 0.0191 0.3755 0.05 0.960
## Year1999 0.6679 0.3701 1.80 0.074 .
## Year2000 0.6689 0.3716 1.80 0.075 .
## Year2001 0.5514 0.3883 1.42 0.159
## Year2002 0.3956 0.3623 1.09 0.278
## Year2003 0.3052 0.3902 0.78 0.436
## Year2004 0.4086 0.4036 1.01 0.314
## Year2005 0.3926 0.3649 1.08 0.285
## Year2006 -0.0466 0.4201 -0.11 0.912
```

```

## Year2007          0.2185      0.3801      0.57      0.567
## Year2008          0.4519      0.3848      1.17      0.243
## Year2009          0.3016      0.3674      0.82      0.414
## Year2010          0.2911      0.4038      0.72      0.473
## Year2011          0.3545      0.4251      0.83      0.406
## Year2012          0.3750      0.3822      0.98      0.329
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.222, Adjusted R-squared:  0.0846
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 103 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.466  0.888  0.960  0.910  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      8.77e-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: 114"
## [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
##   21   12   20   11   13   12   12   13   11   19   20   15   15   25   15
## 2011 2012
##   23   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13    5   10    8    8    5    8    7    7   11    5    7    7    8    6
## 2011 2012

```

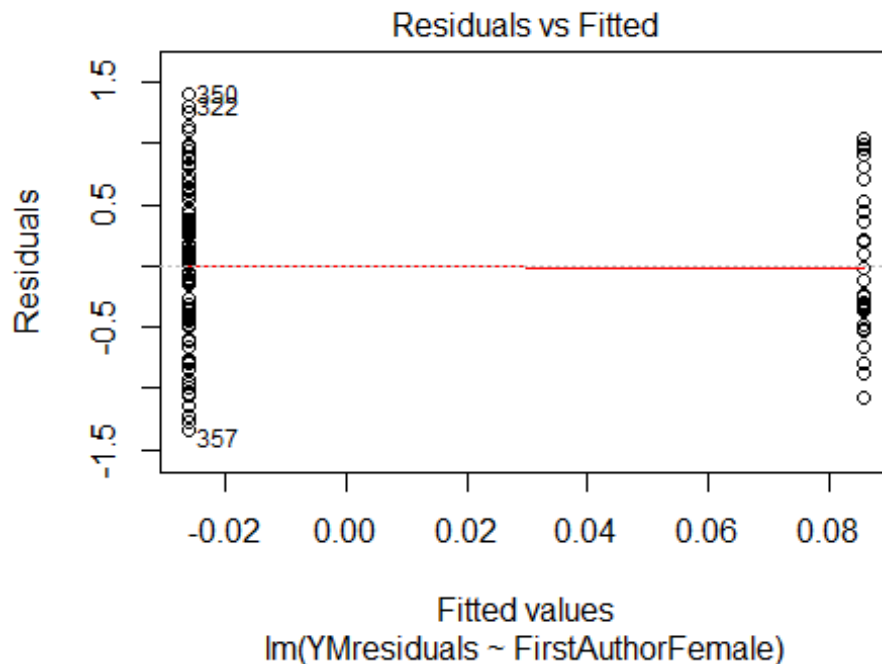
```
##      9      8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13    4    9    8    8    5    7    7    7    8    4    5    6    6    6
## 2011 2012
##    9    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 6.9, df = 16, p-value = 1
```



```
##
## 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"
## [1] "Male first author team size 2018 geometric mean: 1.8612097182042"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

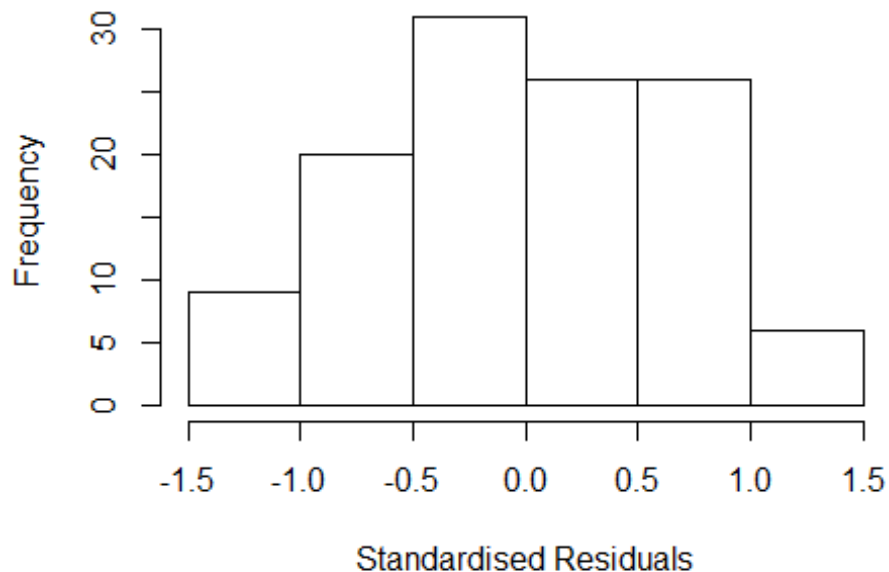
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 1.8612097182042"

## 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 = 5, 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 1.617 1      1.272
## LastAuthorFemale  2.114 1      1.454
## UniqueAuthors    5.890 4      1.248
## Year             11.353 16      1.079
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.44538 -0.48499 -0.00875 0.53808 1.32755
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6256 0.2087 7.79 8.3e-12 ***
## FirstAuthorFemale1 0.2118 0.1610 1.32 0.19143
## LastAuthorFemale1 -0.1923 0.2218 -0.87 0.38816
## UniqueAuthors2 0.0633 0.1826 0.35 0.72961
## UniqueAuthors3 -0.1594 0.1802 -0.88 0.37873
## UniqueAuthors4 -0.0438 0.2726 -0.16 0.87277
## UniqueAuthors5 -0.3047 0.4321 -0.71 0.48233
## Year1997 -0.2435 0.5578 -0.44 0.66343
## Year1998 0.1770 0.2964 0.60 0.55178
## Year1999 -0.8318 0.3055 -2.72 0.00771 **
```

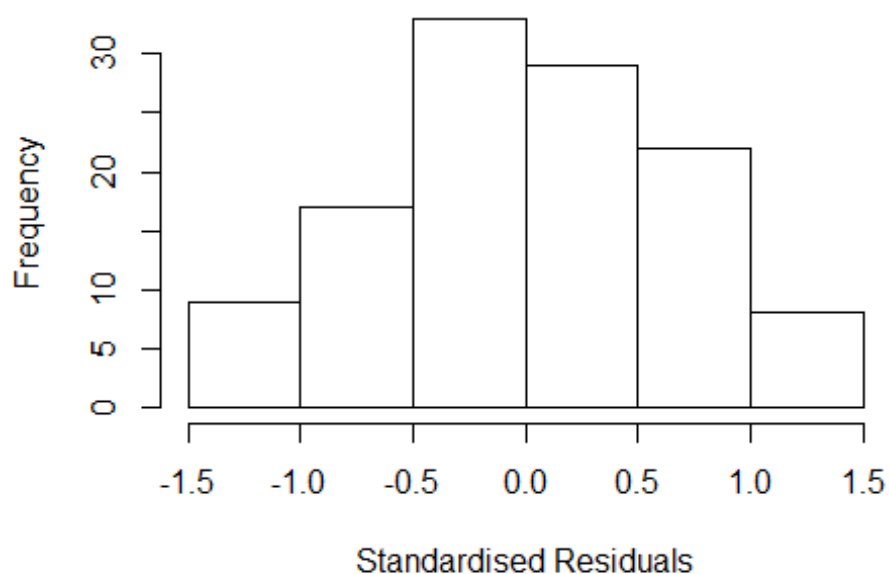


```

## Year2000          -0.8644      0.3449    -2.51  0.01389 *
## Year2001           0.3197      0.4308      0.74  0.45986
## Year2002          -0.2139      0.2608    -0.82  0.41420
## Year2003          -0.1595      0.3515    -0.45  0.65095
## Year2004          -0.2699      0.2498    -1.08  0.28274
## Year2005          -0.5829      0.3976    -1.47  0.14594
## Year2006          -0.8278      0.3199    -2.59  0.01119 *
## Year2007          -0.4791      0.4203    -1.14  0.25725
## Year2008          -0.5227      0.3666    -1.43  0.15719
## Year2009           0.2707      0.2600      1.04  0.30050
## Year2010          -1.2477      0.3662    -3.41  0.00096 ***
## Year2011          -0.5994      0.3443    -1.74  0.08494 .
## Year2012          -0.2181      0.4147    -0.53  0.60011
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.787
## Multiple R-squared:  0.281, Adjusted R-squared:  0.114
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 111 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.717  0.914  0.959  0.936  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      8.47e-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 1.556 1      1.247
## LastAuthorFemale  2.054 1      1.433
## Year              2.454 16      1.028

```

Residuals from first and last author



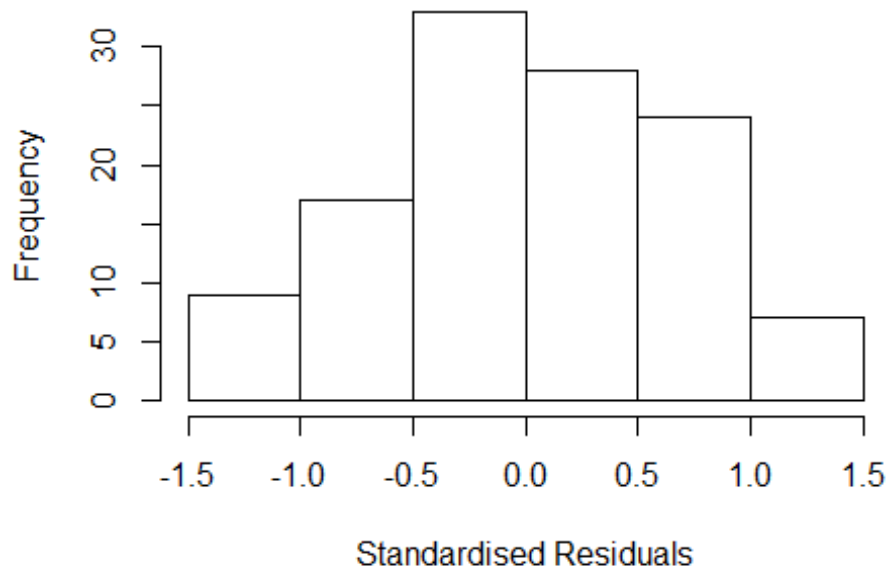
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.42744 -0.46420 -0.00194 0.50106 1.34340
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.584 0.194 8.18 9.7e-13 ***
## FirstAuthorFemale1 0.209 0.164 1.28 0.2042
## LastAuthorFemale1 -0.178 0.223 -0.80 0.4265
## Year1997 -0.170 0.562 -0.30 0.7624
## Year1998 0.223 0.279 0.80 0.4276
## Year1999 -0.813 0.294 -2.77 0.0067 **
## Year2000 -0.818 0.357 -2.29 0.0242 *
## Year2001 0.244 0.410 0.59 0.5537
## Year2002 -0.217 0.266 -0.81 0.4172
## Year2003 -0.105 0.349 -0.30 0.7651
## Year2004 -0.285 0.252 -1.13 0.2601
## Year2005 -0.555 0.366 -1.52 0.1326
```

```

## Year2006          -0.789      0.309   -2.56   0.0121 *
## Year2007          -0.415      0.426   -0.98   0.3315
## Year2008          -0.494      0.382   -1.29   0.1983
## Year2009           0.271      0.284    0.95   0.3421
## Year2010          -1.213      0.360   -3.37   0.0011 **
## Year2011          -0.628      0.360   -1.75   0.0840 .
## Year2012          -0.156      0.384   -0.41   0.6853
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.758
## Multiple R-squared:  0.272, Adjusted R-squared:  0.14
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 112 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.703  0.900  0.961  0.931  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      8.47e-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 1.47 1      1.212
## Year              1.47 16      1.012

```

Residuals from first author



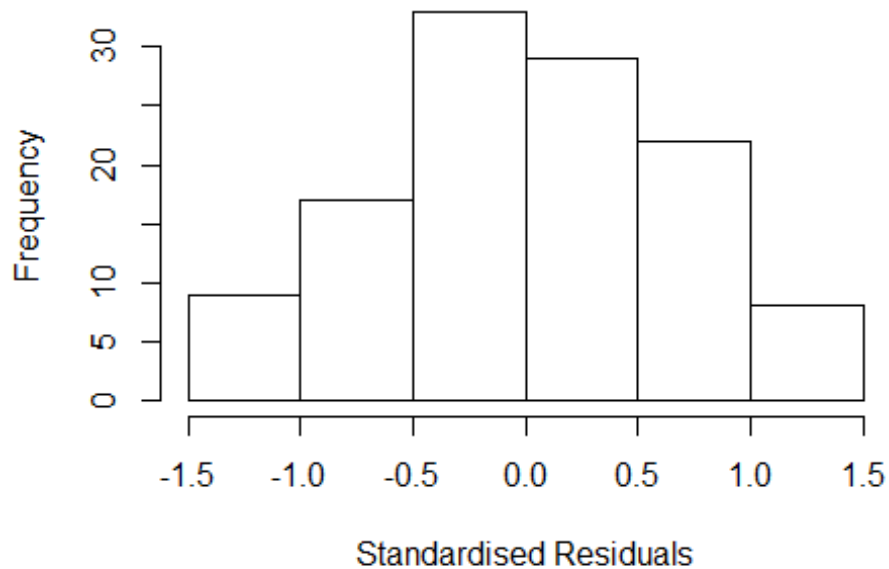
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4455 -0.4416 0.0121 0.5564 1.3782
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5886 0.1974 8.05 1.8e-12 ***
## FirstAuthorFemale1 0.1301 0.1642 0.79 0.42988
## Year1997 -0.1741 0.5650 -0.31 0.75866
## Year1998 0.1981 0.2735 0.72 0.47054
## Year1999 -0.8075 0.2999 -2.69 0.00831 **
## Year2000 -0.8230 0.3587 -2.29 0.02386 *
## Year2001 0.1997 0.4142 0.48 0.63082
## Year2002 -0.2265 0.2624 -0.86 0.39001
## Year2003 -0.0948 0.3514 -0.27 0.78791
## Year2004 -0.3156 0.2473 -1.28 0.20481
## Year2005 -0.5641 0.3726 -1.51 0.13314
## Year2006 -0.8155 0.2968 -2.75 0.00713 **
```

```

## Year2007          -0.4899      0.4246   -1.15   0.25126
## Year2008          -0.5197      0.3868   -1.34   0.18209
## Year2009           0.2778      0.2964    0.94   0.35085
## Year2010          -1.3064      0.3298   -3.96   0.00014 ***
## Year2011          -0.6677      0.3697   -1.81   0.07393 .
## Year2012          -0.1430      0.3828   -0.37   0.70947
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.753
## Multiple R-squared:  0.269, Adjusted R-squared:  0.145
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 112 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.693  0.896  0.955  0.929  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      8.47e-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 1.824 1          1.350
## Year            1.824 16          1.019

```

Residuals from last author



```
## [1] "List of 0 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
## -1.47575 -0.47506 -0.00379 0.54630 1.30506
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6203 0.1935 8.37 3.5e-13 ***
## LastAuthorFemale1 -0.0518 0.2161 -0.24 0.81095
## Year1997 -0.2039 0.5655 -0.36 0.71922
## Year1998 0.2003 0.2752 0.73 0.46835
## Year1999 -0.8233 0.3054 -2.70 0.00824 **
## Year2000 -0.8271 0.3814 -2.17 0.03247 *
## Year2001 0.2409 0.4086 0.59 0.55682
## Year2002 -0.2128 0.2571 -0.83 0.40977
## Year2003 -0.1023 0.3521 -0.29 0.77206
## Year2004 -0.3398 0.2464 -1.38 0.17089
## Year2005 -0.5412 0.3698 -1.46 0.14641
## Year2006 -0.8089 0.3221 -2.51 0.01362 *
```

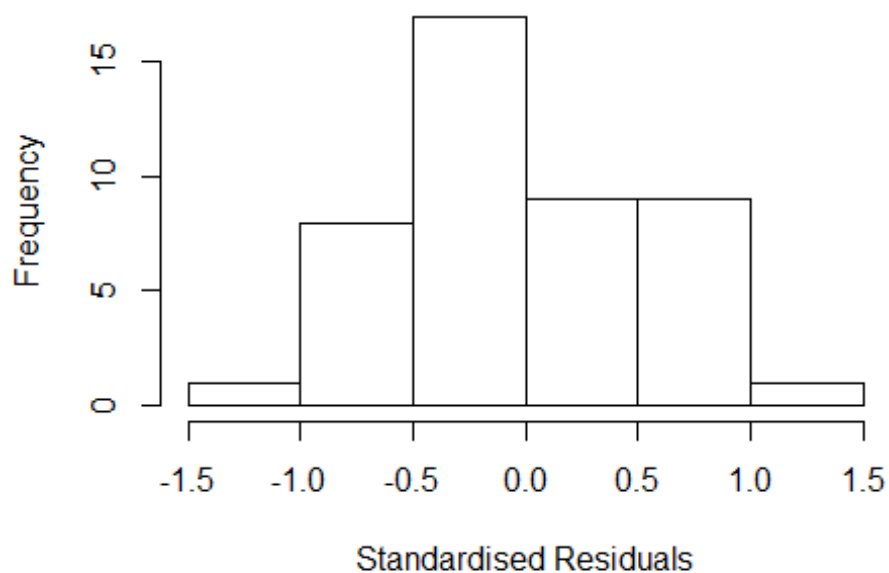
```

## Year2007          -0.5006      0.4274    -1.17   0.24423
## Year2008          -0.5137      0.3804    -1.35   0.17991
## Year2009           0.3180      0.2810     1.13   0.26040
## Year2010          -1.2299      0.3615    -3.40   0.00096 ***
## Year2011          -0.6264      0.3553    -1.76   0.08096 .
## Year2012          -0.1446      0.3773    -0.38   0.70239
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.747
## Multiple R-squared:  0.266, Adjusted R-squared:  0.141
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.676  0.894  0.952  0.926  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      8.47e-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: 118"
## [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 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    1    4    1    7    5    3    3    1    5   11    6    9    2    1
## 2012
##    6
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    2    0    5    3    2    3    1    5    9    6    7    1    1
## 2012

```

```
##      6
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    1    0    4    3    2    2    1    4    9    6    7    1    1
## 2012
##    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 1.73205080756888"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 1.81712059283214"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -6.628e+14  1          NaN
## LastAuthorFemale  8.183e+00  1          2.861
## UniqueAuthors    -2.496e+29  4          NaN
## Year              -1.866e+44 12          NaN
```

Residuals from first and last author and team size




```

## [1] "List of 0 outliers with 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
## -1.23e+00 -3.63e-01 -6.66e-16 3.55e-01 1.21e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5518 0.2392 6.49 7.1e-07 ***
## FirstAuthorFemale1 -0.0308 0.4099 -0.08 0.9407
## LastAuthorFemale1 -0.0274 0.6146 -0.04 0.9648
## UniqueAuthors2 0.1762 0.2392 0.74 0.4680
## UniqueAuthors3 0.3883 0.4269 0.91 0.3713
## UniqueAuthors4 0.0832 0.3582 0.23 0.8181
## UniqueAuthors5 0.4908 0.2641 1.86 0.0744 .
## Year2001 -0.2224 0.2377 -0.94 0.3581
## Year2002 -0.1376 0.3073 -0.45 0.6580
## Year2003 -0.0953 0.7514 -0.13 0.9000
## Year2004 -0.2364 0.4631 -0.51 0.6140
## Year2005 -1.1888 0.2392 -4.97 3.6e-05 ***
## Year2006 -1.0858 0.2781 -3.90 0.0006 ***
## Year2007 -0.8370 0.2976 -2.81 0.0092 **
## Year2008 -0.9105 0.4095 -2.22 0.0351 *
## Year2009 -0.7083 0.2411 -2.94 0.0069 **
## Year2010 0.1180 0.2976 0.40 0.6950
## Year2011 -0.7253 0.2411 -3.01 0.0058 **
## Year2012 -0.7117 0.1928 -3.69 0.0010 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.829
## Multiple R-squared: 0.282, Adjusted R-squared: -0.215
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~1. The remaining 33 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.809 0.936 0.957 0.949 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          2.22e-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 -2.555e+15  1          NaN
## LastAuthorFemale  2.257e+00  1          1.502
## Year              -2.180e+14 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 = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9559 -0.4999  0.0237  0.4127  1.3368
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.73e+00   3.84e-08  4.50e+07 < 2e-16 ***
## FirstAuthorFemale1 -5.99e-02   2.28e-01 -2.60e-01  0.79447
## LastAuthorFemale1 -9.61e-02   3.75e-01 -2.60e-01  0.79937
## Year2001         -2.60e-01   2.47e-01 -1.05e+00  0.30005
## Year2002         -1.86e-01   3.15e-01 -5.90e-01  0.56005
## Year2003         -2.72e-01   7.64e-01 -3.60e-01  0.72479
## Year2004         -3.24e-01   5.53e-01 -5.90e-01  0.56163
## Year2005         -1.36e+00   5.28e-08 -2.59e+07 < 2e-16 ***
## Year2006         -9.65e-01   1.76e-01 -5.48e+00  6e-06 ***
## Year2007         -8.29e-01   2.29e-01 -3.62e+00  0.00107 **
## Year2008         -1.06e+00   4.10e-01 -2.58e+00  0.01498 *
## Year2009         -7.72e-01   2.28e-01 -3.39e+00  0.00197 **
## Year2010          2.50e-02   4.34e-08  5.76e+05 < 2e-16 ***
## Year2011         -7.89e-01   2.28e-01 -3.46e+00  0.00162 **
## Year2012         -6.78e-01   1.84e-01 -3.68e+00  0.00091 ***

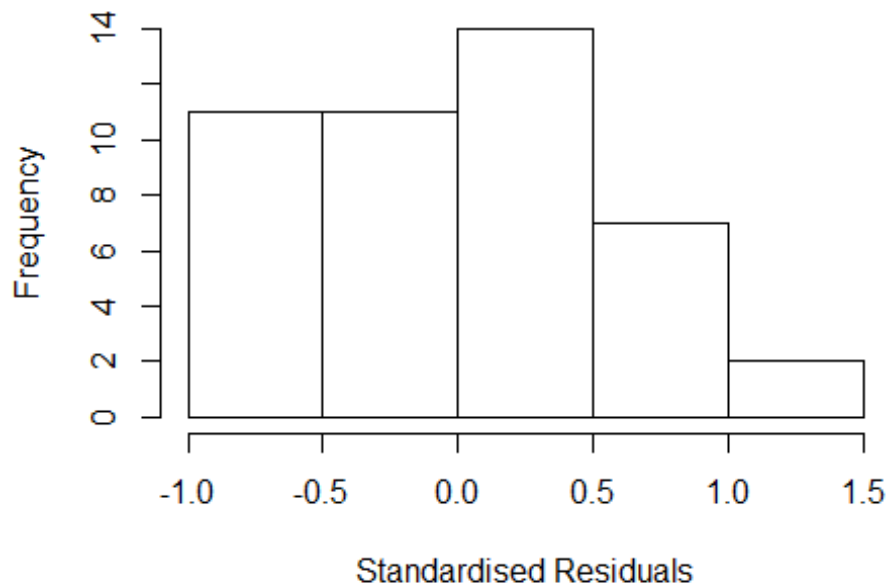
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.738
## Multiple R-squared:  0.243, Adjusted R-squared:  -0.11
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 35 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.723  0.923  0.946  0.936  0.981  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           2.22e-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"
## 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

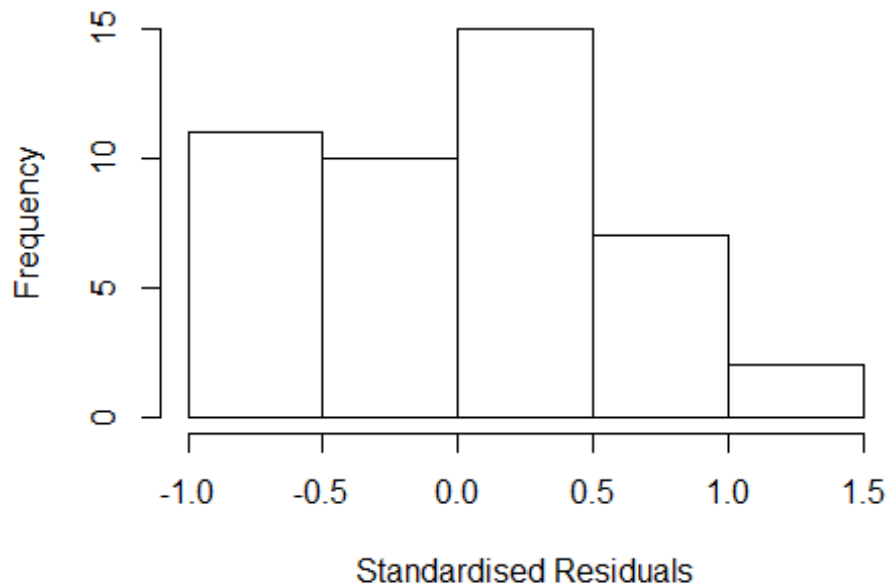
```

Residuals from first and last author



```
##          GVIF Df GVIF^(1/(2*Df))  
## FirstAuthorFemale  NaN  1          NaN  
## Year              NaN 12          NaN
```

Residuals from first author

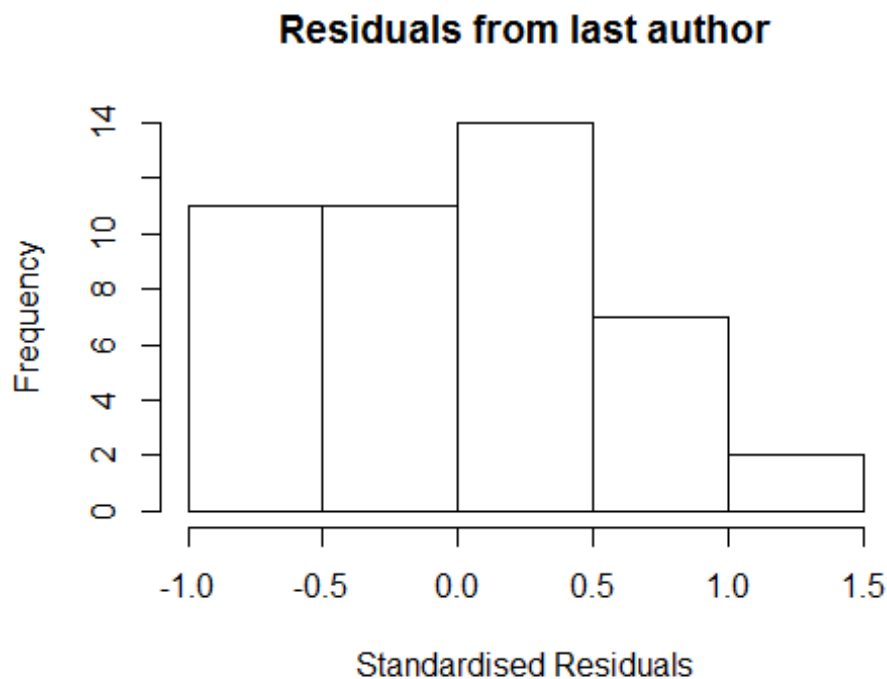


```

## [1] "List of 0 outliers 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
## -0.9681 -0.4999 0.0511 0.4176 1.3385
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.73e+00 3.37e-08 5.13e+07 < 2e-16 ***
## FirstAuthorFemale1 -1.33e-01 2.32e-01 -5.70e-01 0.57013
## Year2001 -2.60e-01 2.47e-01 -1.05e+00 0.30044
## Year2002 -2.00e-01 3.18e-01 -6.30e-01 0.53544
## Year2003 -2.71e-01 7.70e-01 -3.50e-01 0.72671
## Year2004 -3.24e-01 5.55e-01 -5.80e-01 0.56314
## Year2005 -1.36e+00 2.55e-08 -5.36e+07 < 2e-16 ***
## Year2006 -9.70e-01 1.82e-01 -5.32e+00 8.7e-06 ***
## Year2007 -8.34e-01 2.21e-01 -3.78e+00 0.00067 ***
## Year2008 -1.06e+00 4.12e-01 -2.57e+00 0.01507 *
## Year2009 -7.60e-01 2.02e-01 -3.76e+00 0.00072 ***
## Year2010 2.50e-02 0.00e+00 Inf < 2e-16 ***
## Year2011 -7.16e-01 2.32e-01 -3.09e+00 0.00419 **
## Year2012 -6.78e-01 1.84e-01 -3.68e+00 0.00088 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.729
## Multiple R-squared: 0.243, Adjusted R-squared: -0.0738
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 35 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.717 0.920 0.947 0.935 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 2.22e-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 1.403 1           1.185
## Year              1.403 12           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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9473 -0.5000 0.0234 0.4122 1.3425
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.73e+00 4.17e-08 4.15e+07 < 2e-16 ***
```

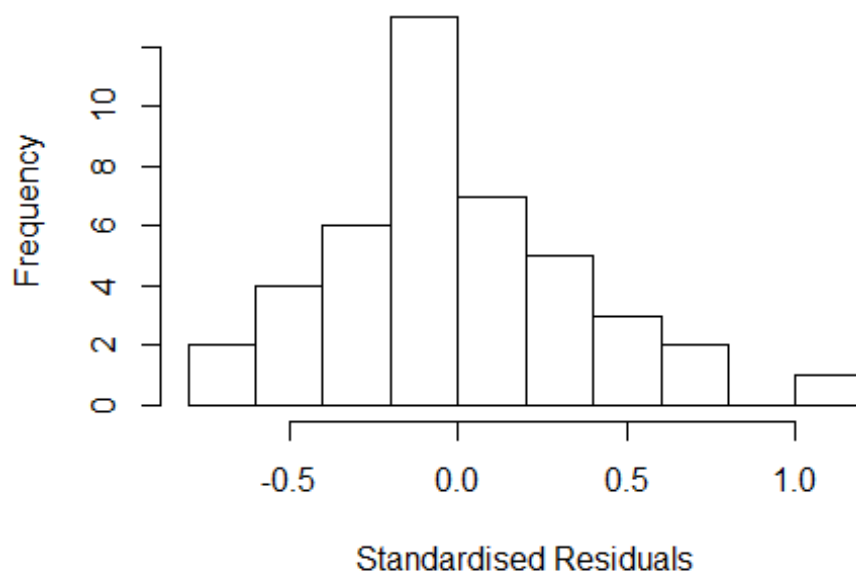
```

## LastAuthorFemale1 -1.56e-01 2.97e-01 -5.20e-01 0.60350
## Year2001 -2.60e-01 2.48e-01 -1.05e+00 0.30201
## Year2002 -1.85e-01 3.16e-01 -5.90e-01 0.56205
## Year2003 -2.71e-01 7.84e-01 -3.50e-01 0.73135
## Year2004 -3.24e-01 5.61e-01 -5.80e-01 0.56697
## Year2005 -1.36e+00 3.33e-08 -4.11e+07 < 2e-16 ***
## Year2006 -9.64e-01 1.76e-01 -5.47e+00 5.6e-06 ***
## Year2007 -8.28e-01 2.30e-01 -3.60e+00 0.00109 **
## Year2008 -1.06e+00 4.18e-01 -2.55e+00 0.01603 *
## Year2009 -7.81e-01 1.91e-01 -4.08e+00 0.00029 ***
## Year2010 2.50e-02 4.53e-08 5.52e+05 < 2e-16 ***
## Year2011 -8.49e-01 4.35e-08 -1.95e+07 < 2e-16 ***
## Year2012 -6.78e-01 1.85e-01 -3.67e+00 0.00089 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.711
## Multiple R-squared: 0.245, Adjusted R-squared: -0.0717
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 35 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.702 0.917 0.940 0.932 0.979 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 2.22e-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: 45"
## [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

```

```
##      2      5      3     13      6      6      3      3      2      2      4      2      6      4     11
## 2011 2012
##      7      6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      2      0      8      4      2      1      2      1      1      1      2      3      2      9
## 2011 2012
##      7      2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      2      0      8      3      2      1      2      1      1      1      2      3      1      9
## 2011 2012
##      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: 1.5874010519682"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale    16.05  1         4.006
## LastAuthorFemale    -18.29  1          NaN
## UniqueAuthors      901.37  3         3.108
## Year               18072.47 14         1.419
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
```



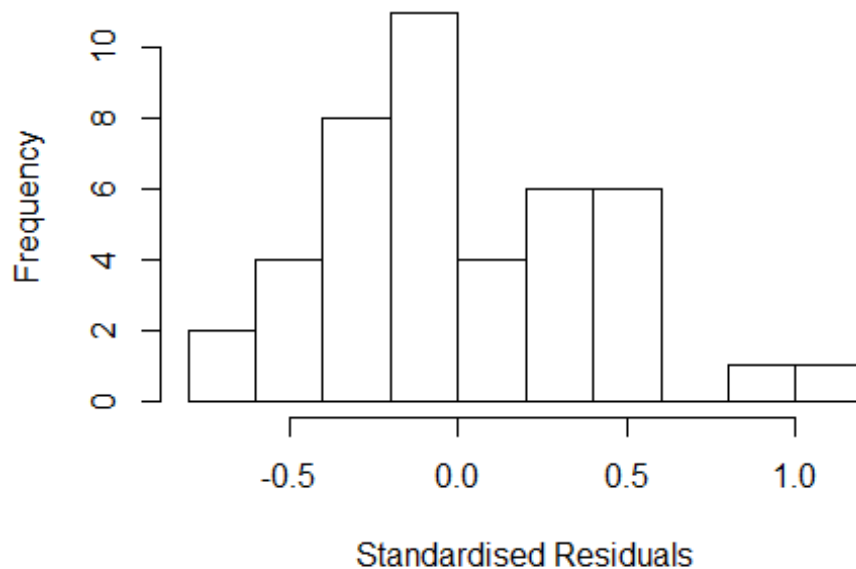
```

## <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
## -7.80e-01 -2.66e-01 -2.22e-16  2.42e-01  1.05e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1045     0.2590   4.26 0.00029 ***
## FirstAuthorFemale1  0.0543     0.4116   0.13 0.89613
## LastAuthorFemale1 -0.3976     0.4467  -0.89 0.38264
## UniqueAuthors2     0.2733     0.2502   1.09 0.28594
## UniqueAuthors3     0.4458     0.4048   1.10 0.28221
## UniqueAuthors4     0.3949     0.4968   0.80 0.43473
## Year1999           -0.0231     0.2214  -0.10 0.91776
## Year2000           -0.5687     0.3042  -1.87 0.07434 .
## Year2001           -0.8899     0.6713  -1.33 0.19797
## Year2002           -0.7475     0.2590  -2.89 0.00834 **
## Year2003           -0.5902     0.3596  -1.64 0.11435
## Year2004           -1.1045     0.2590  -4.26 0.00029 ***
## Year2005            0.2685     0.2590   1.04 0.31062
## Year2006            0.4375     0.2590   1.69 0.10468
## Year2007            0.4046     0.5881   0.69 0.49833
## Year2008           -0.3729     0.7642  -0.49 0.63018
## Year2009           -0.4845     0.2590  -1.87 0.07420 .
## Year2010           -0.0938     0.2798  -0.34 0.74056
## Year2011           -0.6232     0.3274  -1.90 0.06956 .
## Year2012           -0.0513     0.2862  -0.18 0.85920
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.508
## Multiple R-squared:  0.464, Adjusted R-squared:  0.0218
## Convergence in 35 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~ = 1. The remaining 36 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.650  0.940  0.959  0.942  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          2.33e-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 0.447 1 0.6686
## LastAuthorFemale 30.377 1 5.5116
## Year 2.011 14 1.0253
```

Residuals from first and last author



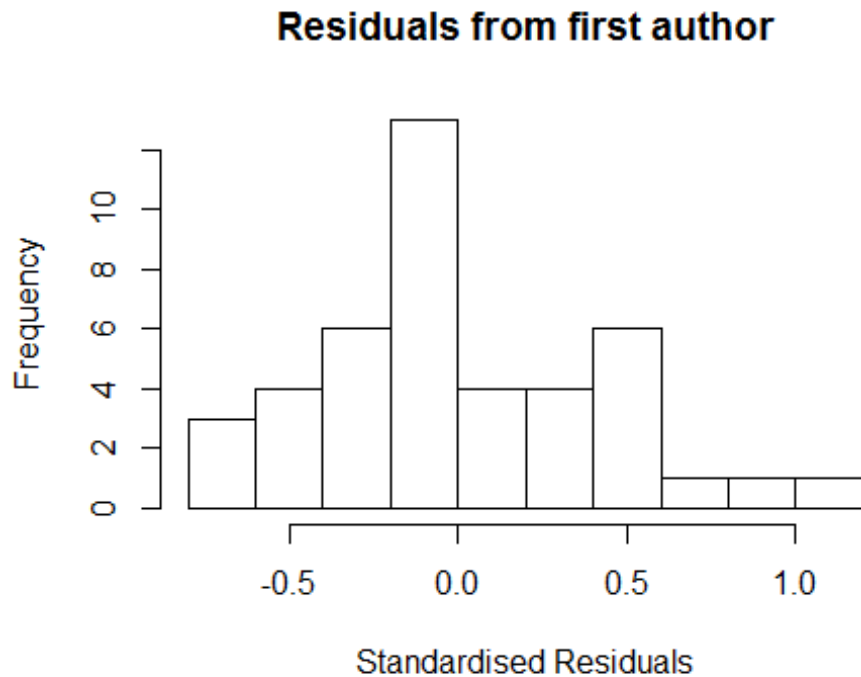
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
```

```

## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -6.77e-01 -2.78e-01  1.11e-16  3.02e-01  1.10e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2622     0.1286   9.81 3.1e-10 ***
## FirstAuthorFemale1 -0.1454     0.2626  -0.55  0.5845
## LastAuthorFemale1  -0.1665     0.2434  -0.68  0.5001
## Year1999          -0.0822     0.2017  -0.41  0.6868
## Year2000          -0.4880     0.1741  -2.80  0.0094 **
## Year2001          -0.8502     0.3440  -2.47  0.0203 *
## Year2002          -0.9052     0.1286  -7.04 1.8e-07 ***
## Year2003          -0.7426     0.2982  -2.49  0.0195 *
## Year2004          -1.2622     0.1286  -9.81 3.1e-10 ***
## Year2005           0.1108     0.1286   0.86  0.3971
## Year2006           0.2798     0.1286   2.17  0.0389 *
## Year2007           0.2312     0.5391   0.43  0.6716
## Year2008          -0.4186     0.5960  -0.70  0.4887
## Year2009          -0.6422     0.1286  -4.99 3.4e-05 ***
## Year2010          -0.1136     0.1731  -0.66  0.5175
## Year2011          -0.4414     0.2234  -1.98  0.0589 .
## Year2012          -0.1880     0.2351  -0.80  0.4312
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.576
## Multiple R-squared:  0.405, Adjusted R-squared:  0.0396
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 34 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.695  0.937  0.966  0.947  0.981  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      2.33e-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 -0.3097 1          NaN
## Year              -0.3097 14         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
## -7.72e-01 -2.62e-01 -2.22e-16  3.14e-01  1.01e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.17900    0.10150   11.62  5.2e-12 ***
## FirstAuthorFemale1 -0.25029    0.21988   -1.14  0.2650
## Year1999        -0.00802    0.19906   -0.04  0.9682
## Year2000        -0.40476    0.15441   -2.62  0.0142 *
## Year2001        -0.76700    0.33660   -2.28  0.0308 *
## Year2002        -0.82200    0.10150   -8.10  1.1e-08 ***
```

```

## Year2003          -0.77336      0.34871      -2.22      0.0352 *
## Year2004          -1.17900      0.10150     -11.62     5.2e-12 ***
## Year2005           0.19400      0.10150       1.91      0.0666 .
## Year2006           0.36300      0.10150       3.58      0.0013 **
## Year2007           0.28364      0.58086       0.49      0.6293
## Year2008          -0.40748      0.56159      -0.73      0.4743
## Year2009          -0.55900      0.10150     -5.51     7.8e-06 ***
## Year2010          -0.05049      0.17156      -0.29      0.7708
## Year2011          -0.34766      0.18959     -1.83      0.0777 .
## Year2012          -0.18800      0.31983     -0.59      0.5615
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.586
## Multiple R-squared:  0.394, Adjusted R-squared:  0.057
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
##  6 weights are ~= 1. The remaining 37 ones are summarized as
##    Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##    0.750  0.940  0.963  0.952  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      2.33e-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 43.5  1          6.595
## Year            43.5 14          1.144
##
## [1] "List of 0 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
## -6.33e-01 -2.78e-01  2.22e-16  2.94e-01  1.15e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2936    0.1115   11.60  5.3e-12 ***
## LastAuthorFemale1 -0.2292    0.2196   -1.04  0.3058
## Year1999         -0.1255    0.1650   -0.76  0.4535
## Year2000         -0.5195    0.1620   -3.21  0.0034 **
## Year2001         -0.8816    0.3407   -2.59  0.0154 *
## Year2002         -0.9366    0.1115   -8.40  5.2e-09 ***
## Year2003         -0.7839    0.2361   -3.32  0.0026 **
## Year2004         -1.2936    0.1115  -11.60  5.3e-12 ***
## Year2005          0.0794    0.1115    0.71  0.4825
## Year2006          0.2484    0.1115    2.23  0.0344 *
## Year2007          0.1585    0.5979    0.27  0.7930
## Year2008         -0.4336    0.6229   -0.70  0.4923
## Year2009         -0.6736    0.1115   -6.04  1.9e-06 ***
## Year2010         -0.1384    0.1720   -0.80  0.4280
## Year2011         -0.5172    0.1859   -2.78  0.0097 **
## Year2012         -0.1880    0.2073   -0.91  0.3725
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.556
## Multiple R-squared:  0.406, Adjusted R-squared:  0.0756
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
##  9 weights are ~= 1. The remaining 34 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.650  0.933  0.960   0.943  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      2.33e-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: 43"
## [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 2004 2005 2006 2007 2008 2009 2010 2011
##    2    3    2   11    2    2    1    2    2    2    2    2    3    2    7
##
## 1996 1997 1998 1999 2000 2001 2002 2004 2005 2006 2007 2008 2009 2010 2011
##    0    1    0    6    2    2    0    2    1    2    1    1    2    1    4
##
## 1996 1997 1998 1999 2000 2001 2002 2004 2005 2006 2007 2008 2009 2010 2011
##    0    1    0    6    2    2    0    2    1    2    1    1    2    1    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1"
## [1] "Male first author team size 2018 geometric mean: 4"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 4"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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"

## 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"
## [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: 23"
## [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
##    5    8    6    5    9    6    3   13   13   16   16   18   33   24   22
## 2011 2012
##   31   29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    3    5    2    1    9    5    8    2   12   20   12   14
## 2011 2012
##   12   18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    1    3    2    1    7    3    4    2   12   17   12   12
## 2011 2012
##   10   16
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.64398437900346"
## [1] "Male first author team size 2018 geometric mean: 4.52323980326712"

## 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 = 39, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.16014057648204"
## [1] "Male last author team size 2018 geometric mean: 4.18032194912137"

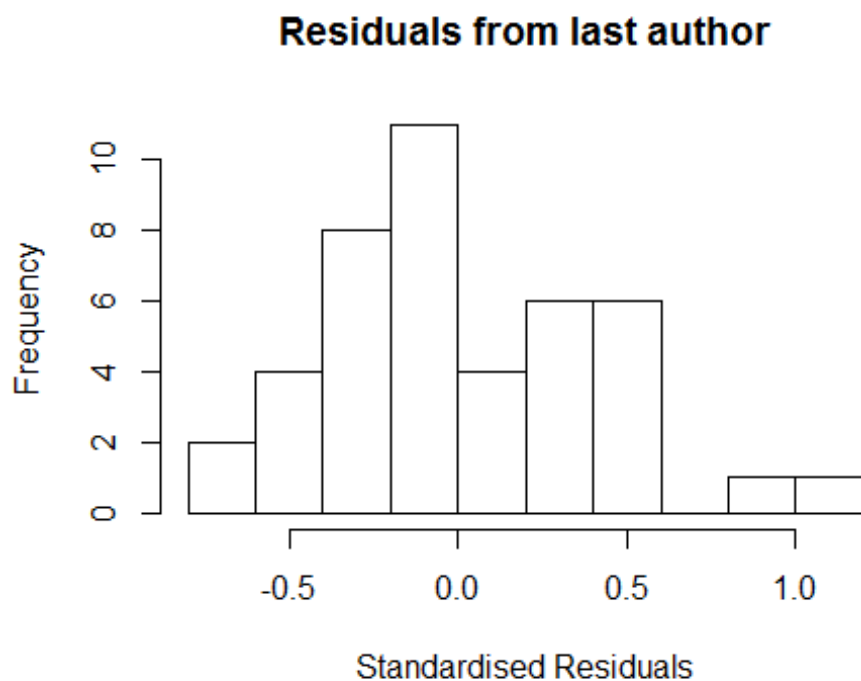
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 46, 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"

```

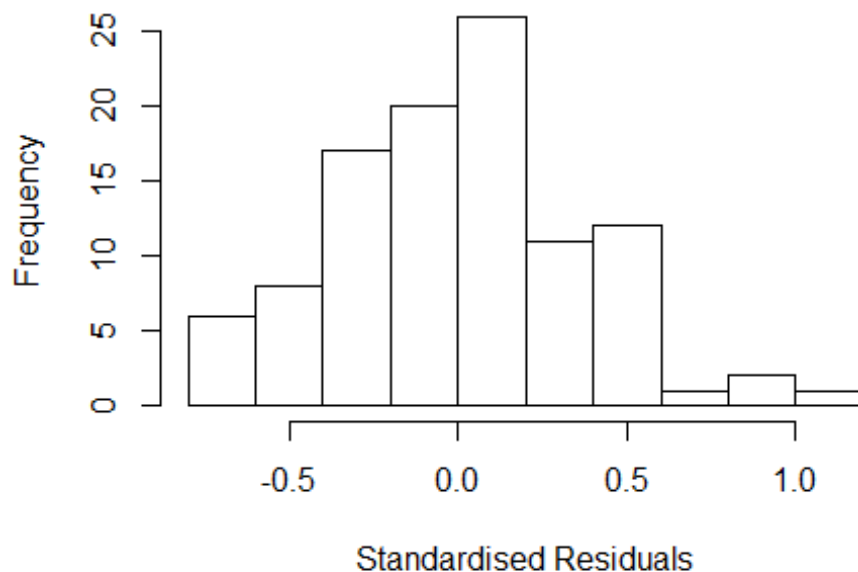


```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```



##		GVIF	Df	GVIF ^{1/(2*Df)}
##	FirstAuthorFemale	-6.257e+13	1	NaN
##	LastAuthorFemale	2.559e+00	1	1.60
##	UniqueAuthors	5.171e+30	4	6905.59
##	Year	2.990e+44	15	30.38

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -0.78888 -0.21455 0.00243 0.21203 1.09209
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6160 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 0.0623 0.1107 0.56 0.57512
## LastAuthorFemale1 -0.3708 0.2273 -1.63 0.10659
## UniqueAuthors2 0.2844 0.4006 0.71 0.47968
## UniqueAuthors3 0.3551 0.3820 0.93 0.35533
## UniqueAuthors4 0.3516 0.3906 0.90 0.37060
## UniqueAuthors5 0.3143 0.3885 0.81 0.42087
## Year1998 -1.1516 0.3906 -2.95 0.00416 **
## Year1999 -0.9040 0.4064 -2.22 0.02889 *
## Year2000 -1.5954 0.4273 -3.73 0.00035 ***
```

```

## Year2001      -0.7584      0.3872      -1.96      0.05357 .
## Year2002      -0.8974      0.4006      -2.24      0.02778 *
## Year2003      -0.6400      0.3272      -1.96      0.05385 .
## Year2004      -0.8391      0.4217      -1.99      0.04993 *
## Year2005      -0.6699      0.4000      -1.67      0.09781 .
## Year2006      -0.9716      0.3872      -2.51      0.01407 *
## Year2007      -0.8948      0.4061      -2.20      0.03039 *
## Year2008      -0.8484      0.3962      -2.14      0.03519 *
## Year2009      -0.9196      0.3987      -2.31      0.02359 *
## Year2010      -0.7368      0.3998      -1.84      0.06894 .
## Year2011      -0.8192      0.4345      -1.89      0.06293 .
## Year2012      -0.8669      0.4224      -2.05      0.04335 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.211, Adjusted R-squared:  0.00844
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 89 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.375  0.863   0.938   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           9.62e-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 1.593e+16 1 1.262e+08
## LastAuthorFemale 2.330e+00 1 1.526e+00
## Year 3.366e+16 15 3.556e+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
## -1.010 -0.200  0.001  0.193  1.093
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.62e+00   3.77e-08  4.29e+07 < 2e-16 ***
## FirstAuthorFemale1 6.59e-02   1.10e-01  6.00e-01  0.55097
## LastAuthorFemale1 -3.59e-01   2.20e-01 -1.63e+00  0.10582
## Year1998         -8.00e-01   4.25e-08 -1.88e+07 < 2e-16 ***
## Year1999         -5.56e-01   1.10e-01 -5.05e+00  2.5e-06 ***
## Year2000         -1.26e+00   1.87e-01 -6.76e+00  1.6e-09 ***
## Year2001         -4.59e-01   1.41e-03 -3.25e+02 < 2e-16 ***
## Year2002         -6.13e-01   4.44e-08 -1.38e+07 < 2e-16 ***
## Year2003         -3.56e-01   9.29e-02 -3.83e+00  0.00024 ***
## Year2004         -5.00e-01   1.85e-01 -2.70e+00  0.00837 **
## Year2005         -3.49e-01   1.30e-01 -2.68e+00  0.00882 **
## Year2006         -6.17e-01   6.32e-02 -9.76e+00  1.4e-15 ***
## Year2007         -6.06e-01   1.71e-01 -3.54e+00  0.00065 ***
## Year2008         -5.26e-01   9.09e-02 -5.78e+00  1.2e-07 ***
## Year2009         -6.01e-01   8.23e-02 -7.30e+00  1.3e-10 ***
## Year2010         -4.13e-01   8.56e-02 -4.82e+00  6.1e-06 ***
## Year2011         -4.83e-01   2.01e-01 -2.41e+00  0.01811 *
## Year2012         -5.28e-01   1.55e-01 -3.42e+00  0.00097 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.193, Adjusted R-squared:  0.0339
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 88 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.400  0.879  0.941  0.903  0.988  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      9.62e-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

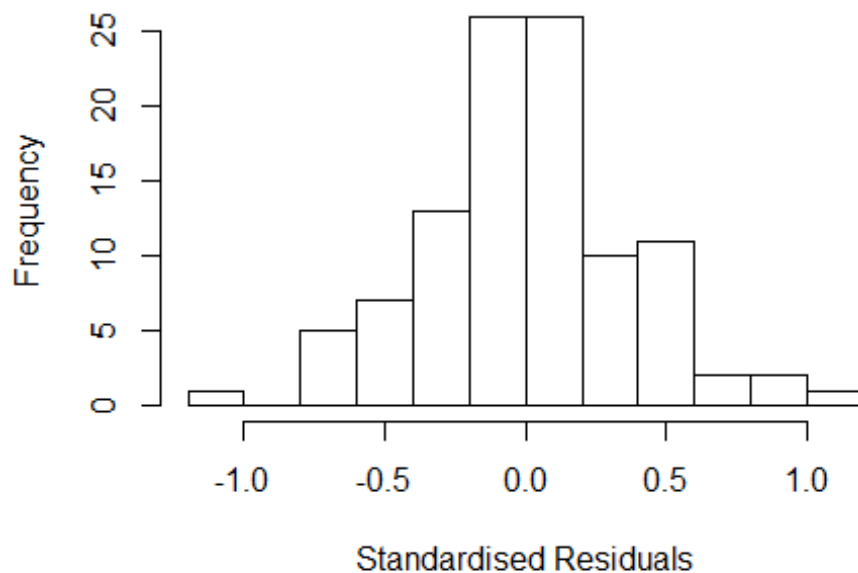
```

```
##           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"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

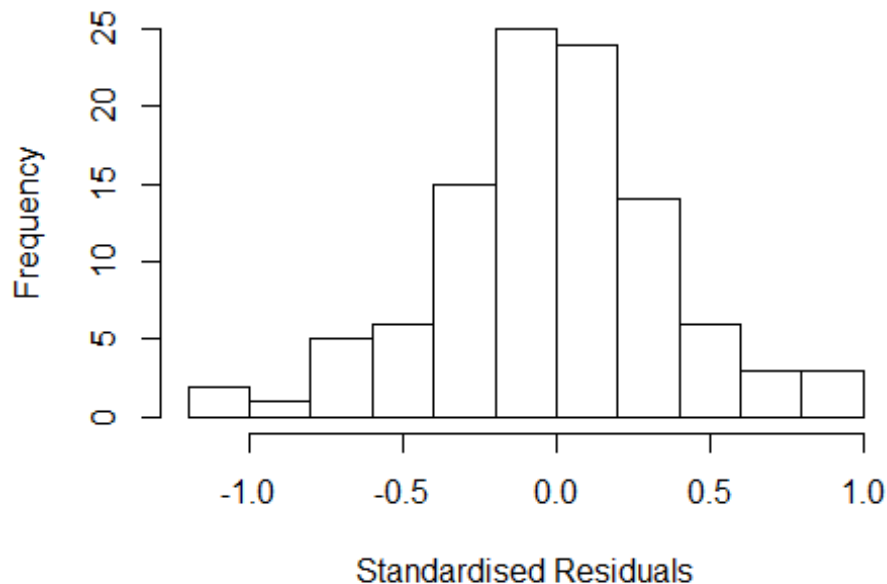
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1 NaN
## Year NaN 15 NaN
```

Residuals from first author



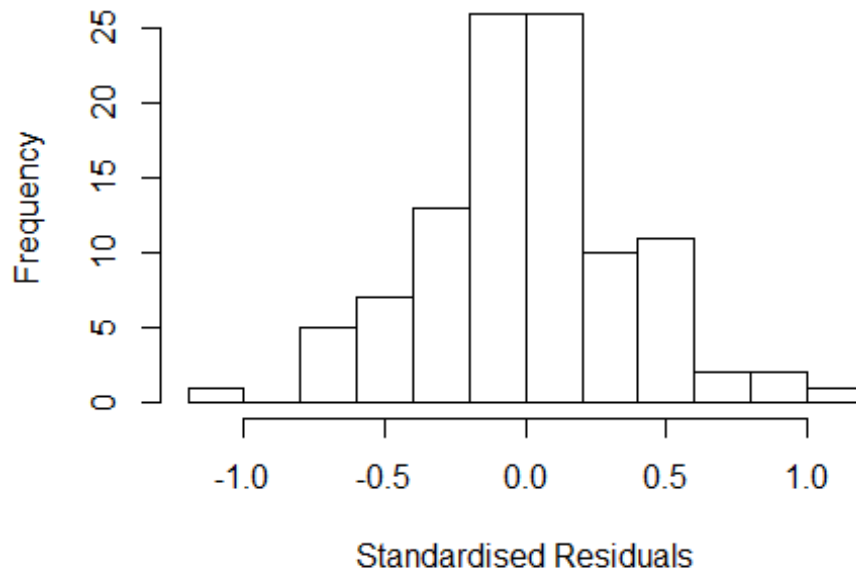
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.08e+00 -2.37e-01  2.78e-16  1.98e-01  9.68e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.62e+00   2.97e-08  5.44e+07 < 2e-16 ***
## FirstAuthorFemale1 -2.22e-02   1.13e-01 -2.00e-01  0.84472
## Year1998        -8.00e-01   0.00e+00    -Inf < 2e-16 ***
## Year1999        -4.68e-01   1.13e-01 -4.15e+00  7.8e-05 ***
## Year2000        -1.26e+00   1.85e-01 -6.80e+00  1.3e-09 ***
## Year2001        -4.59e-01   1.41e-03 -3.25e+02 < 2e-16 ***
## Year2002        -6.13e-01   0.00e+00    -Inf < 2e-16 ***
## Year2003        -3.94e-01   1.20e-01 -3.27e+00  0.00152 **
## Year2004        -4.67e-01   1.72e-01 -2.71e+00  0.00810 **
## Year2005        -5.02e-01   1.46e-01 -3.45e+00  0.00087 ***
## Year2006        -6.17e-01   6.31e-02 -9.76e+00  1.2e-15 ***
## Year2007        -6.18e-01   1.41e-01 -4.37e+00  3.4e-05 ***
```

```

## Year2008          -5.16e-01    8.92e-02 -5.79e+00    1.1e-07 ***
## Year2009          -5.85e-01    8.18e-02 -7.15e+00    2.5e-10 ***
## Year2010          -4.31e-01    8.89e-02 -4.84e+00    5.5e-06 ***
## Year2011          -5.26e-01    2.15e-01 -2.45e+00    0.01628 *
## Year2012          -5.40e-01    1.60e-01 -3.37e+00    0.00113 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.14,   Adjusted R-squared:  -0.0187
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 92 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.445  0.885   0.959   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          9.62e-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 2.423 1          1.557
## Year            2.423 15          1.030

```

Residuals from last author



```
## [1] "List of 0 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
## -1.02106 -0.22039  0.00203  0.19429  1.11455
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.62e+00  4.07e-08  3.97e+07 < 2e-16 ***
## LastAuthorFemale1 -3.25e-01  2.11e-01 -1.54e+00  0.12762
## Year1998        -8.00e-01  2.50e-08 -3.20e+07 < 2e-16 ***
## Year1999        -4.90e-01  4.34e-08 -1.13e+07 < 2e-16 ***
## Year2000        -1.26e+00  1.87e-01 -6.74e+00  1.7e-09 ***
## Year2001        -4.59e-01  1.41e-03 -3.25e+02 < 2e-16 ***
## Year2002        -6.13e-01  2.93e-08 -2.09e+07 < 2e-16 ***
## Year2003        -3.54e-01  1.01e-01 -3.51e+00  0.00071 ***
## Year2004        -4.76e-01  1.68e-01 -2.83e+00  0.00581 **
## Year2005        -3.50e-01  1.26e-01 -2.77e+00  0.00678 **
## Year2006        -6.17e-01  6.32e-02 -9.75e+00  1.3e-15 ***
## Year2007        -5.95e-01  1.59e-01 -3.74e+00  0.00033 ***
```

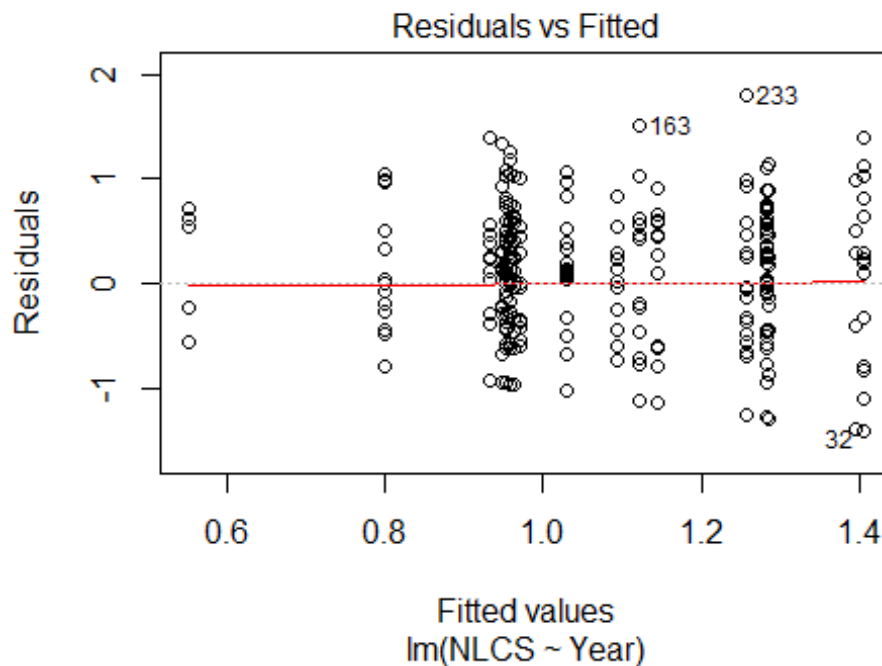


```

## Year2008          -5.05e-01    8.97e-02 -5.63e+00    2.1e-07 ***
## Year2009          -5.89e-01    7.80e-02 -7.55e+00    4.0e-11 ***
## Year2010          -4.03e-01    8.26e-02 -4.87e+00    4.9e-06 ***
## Year2011          -4.73e-01    2.10e-01 -2.26e+00    0.02640 *
## Year2012          -5.20e-01    1.57e-01 -3.32e+00    0.00131 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.378
## Multiple R-squared:  0.192, Adjusted R-squared:  0.0432
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 92 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.366  0.866  0.953  0.904  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      9.62e-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: 104"
## [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
##   24   37   23   31   17   36   35   20   26   23   21   30   37   27   52
## 2011 2012
##   59   52
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5   15   10   14    7   13   20   11   15   13   11   21   21   13   40
## 2011 2012
##   34   31

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5   13   10   13    7   13   18   11   14   12   10   16   20   12   38
## 2011 2012
##   31   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 = 11, df = 16, p-value = 0.8
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.48, df = 1, p-value = 0.5

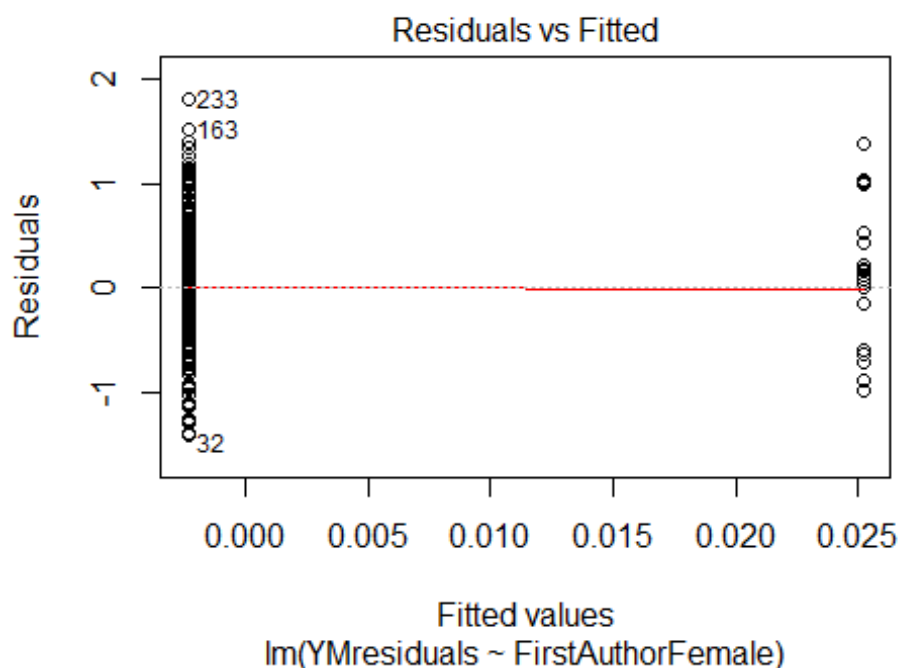
## [1] "Female first author team size 2018 geometric mean: 3.06439349308142"
## [1] "Male first author team size 2018 geometric mean: 2.68394060688961"

## 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 = 90, p-value = 0.6
## 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: 2.85965490078036"

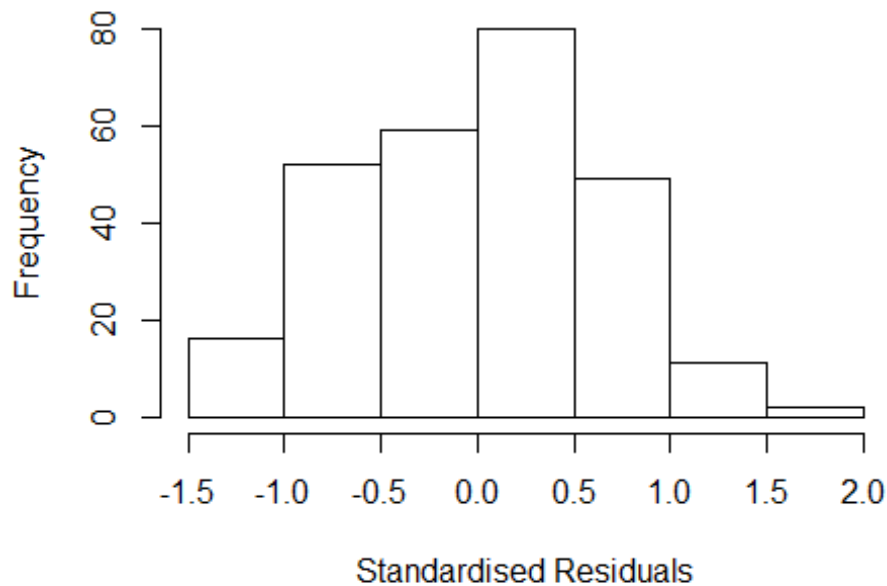
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22, 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	1.655	1	1.286
LastAuthorFemale	1.619	1	1.272
UniqueAuthors	1.968	4	1.088
Year	2.657	16	1.031

Residuals from first and last author and team size



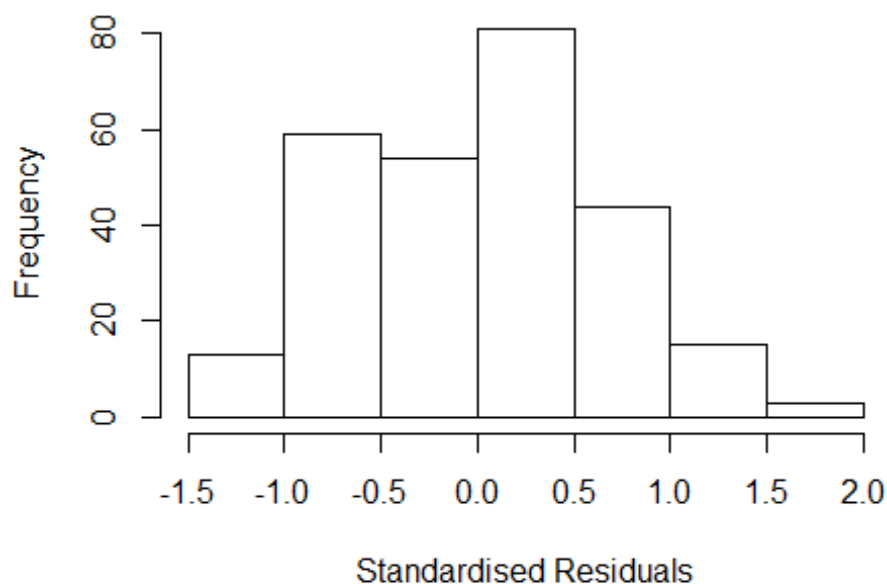
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3773 -0.5041 0.0494 0.4568 1.7897
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.37726 0.43234 3.19 0.0016 **
## FirstAuthorFemale1 -0.00988 0.15821 -0.06 0.9502
## LastAuthorFemale1 -0.01229 0.17014 -0.07 0.9425
## UniqueAuthors2 0.23175 0.10831 2.14 0.0334 *
## UniqueAuthors3 0.02899 0.11801 0.25 0.8062
## UniqueAuthors4 0.59270 0.14703 4.03 7.4e-05 ***
## UniqueAuthors5 0.43953 0.20181 2.18 0.0304 *
## Year1997 -0.24051 0.53002 -0.45 0.6504
## Year1998 -0.12636 0.49103 -0.26 0.7971
## Year1999 -0.31383 0.47526 -0.66 0.5097
```

```

## Year2000      -0.87314      0.48366      -1.81      0.0723 .
## Year2001      -0.43403      0.49720      -0.87      0.3835
## Year2002      -0.34067      0.46200      -0.74      0.4616
## Year2003      -0.50337      0.45280      -1.11      0.2674
## Year2004      -0.71061      0.45879      -1.55      0.1227
## Year2005      -0.64838      0.45350      -1.43      0.1541
## Year2006      -0.25136      0.45111      -0.56      0.5779
## Year2007      -0.56077      0.46059      -1.22      0.2246
## Year2008      -0.58859      0.45894      -1.28      0.2009
## Year2009      -0.73211      0.45745      -1.60      0.1108
## Year2010      -0.28306      0.44396      -0.64      0.5243
## Year2011      -0.58799      0.45097      -1.30      0.1935
## Year2012      -0.63630      0.45380      -1.40      0.1621
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.686
## Multiple R-squared:  0.139, Adjusted R-squared:  0.0623
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 249 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.476  0.894  0.951  0.923  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          3.72e-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 1.697 1          1.303
## LastAuthorFemale  1.690 1          1.300
## Year              1.422 16          1.011

```

Residuals from first and last author



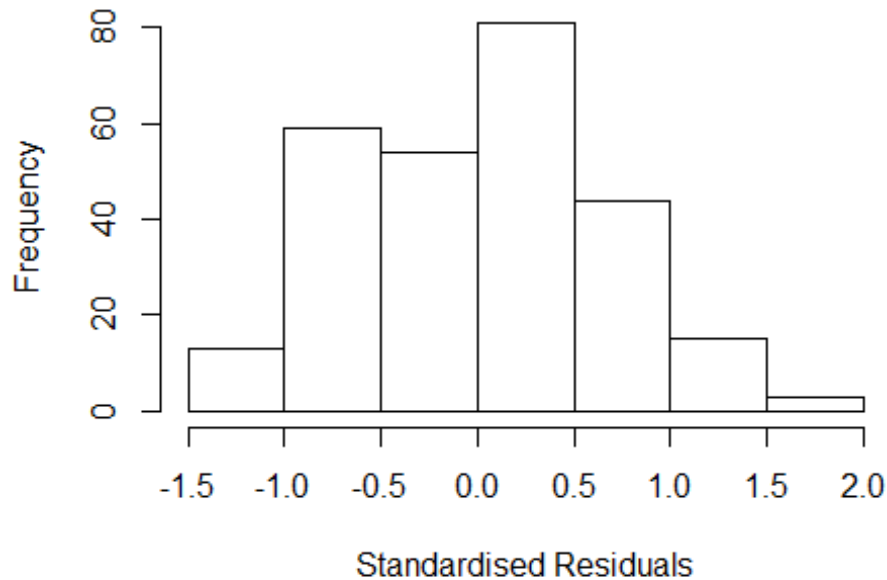
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4677 -0.5439 0.0566 0.4678 1.8833
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.46770 0.41827 3.51 0.00053 ***
## FirstAuthorFemale1 0.00178 0.18729 0.01 0.99241
## LastAuthorFemale1 0.00208 0.20663 0.01 0.99196
## Year1997 -0.17493 0.52168 -0.34 0.73767
## Year1998 -0.17264 0.48663 -0.35 0.72306
## Year1999 -0.23015 0.46557 -0.49 0.62150
## Year2000 -0.92375 0.47819 -1.93 0.05452 .
## Year2001 -0.36749 0.49299 -0.75 0.45670
## Year2002 -0.29305 0.44792 -0.65 0.51356
## Year2003 -0.50953 0.44719 -1.14 0.25563
## Year2004 -0.62571 0.45475 -1.38 0.17007
## Year2005 -0.62975 0.44668 -1.41 0.15983
```

```

## Year2006      -0.31234      0.44141      -0.71      0.47985
## Year2007      -0.52225      0.44877      -1.16      0.24563
## Year2008      -0.44161      0.44258      -1.00      0.31934
## Year2009      -0.63545      0.44911      -1.41      0.15833
## Year2010      -0.16986      0.43214      -0.39      0.69460
## Year2011      -0.50707      0.43762      -1.16      0.24769
## Year2012      -0.53954      0.44032      -1.23      0.22160
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.723
## Multiple R-squared:  0.0749, Adjusted R-squared:  0.00833
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 245 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.478  0.883   0.949   0.924   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      3.72e-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 1.176 1      1.084
## Year      1.176 16      1.005

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.4683 -0.5439 0.0564 0.4679 1.8836
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.46827 0.41850 3.51 0.00053 ***
## FirstAuthorFemale1 0.00269 0.15593 0.02 0.98626
## Year1997 -0.17522 0.52553 -0.33 0.73910
## Year1998 -0.17289 0.48618 -0.36 0.72244
## Year1999 -0.23029 0.46578 -0.49 0.62144
## Year2000 -0.92437 0.48130 -1.92 0.05592 .
## Year2001 -0.36826 0.49285 -0.75 0.45564
## Year2002 -0.29383 0.44790 -0.66 0.51241
## Year2003 -0.51018 0.44739 -1.14 0.25522
## Year2004 -0.62628 0.45466 -1.38 0.16960
## Year2005 -0.63024 0.44710 -1.41 0.15989
## Year2006 -0.31290 0.44163 -0.71 0.47929
```

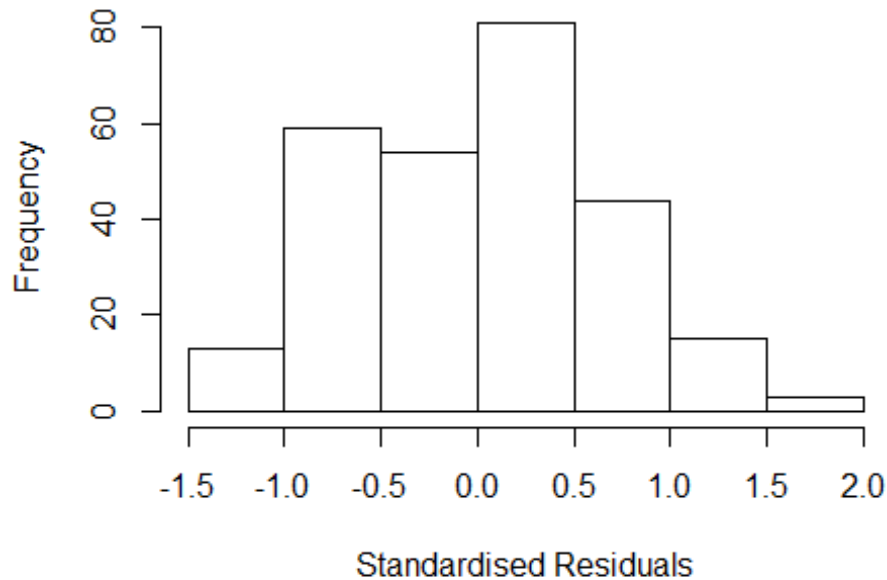


```

## Year2007      -0.52269    0.44873   -1.16  0.24520
## Year2008      -0.44234    0.44243   -1.00  0.31837
## Year2009      -0.63592    0.44937   -1.42  0.15827
## Year2010      -0.17012    0.43194   -0.39  0.69403
## Year2011      -0.50742    0.43582   -1.16  0.24541
## Year2012      -0.54004    0.44076   -1.23  0.22163
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.721
## Multiple R-squared:  0.075, Adjusted R-squared:  0.0124
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 245 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.475  0.882  0.949  0.924  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      3.72e-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 1.173 1      1.083
## Year      1.173 16      1.005

```

Residuals from last author



```
## [1] "List of 0 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
## -1.4685 -0.5439 0.0564 0.4680 1.8837
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.46851 0.41874 3.51 0.00054 ***
## LastAuthorFemale1 0.00276 0.17283 0.02 0.98726
## Year1997 -0.17584 0.52248 -0.34 0.73673
## Year1998 -0.17312 0.48679 -0.36 0.72242
## Year1999 -0.23038 0.46577 -0.49 0.62130
## Year2000 -0.92464 0.48165 -1.92 0.05602 .
## Year2001 -0.36843 0.49363 -0.75 0.45614
## Year2002 -0.29421 0.44837 -0.66 0.51231
## Year2003 -0.51047 0.44769 -1.14 0.25528
## Year2004 -0.62657 0.45527 -1.38 0.16997
## Year2005 -0.63049 0.44716 -1.41 0.15978
## Year2006 -0.31315 0.44186 -0.71 0.47917
```

```

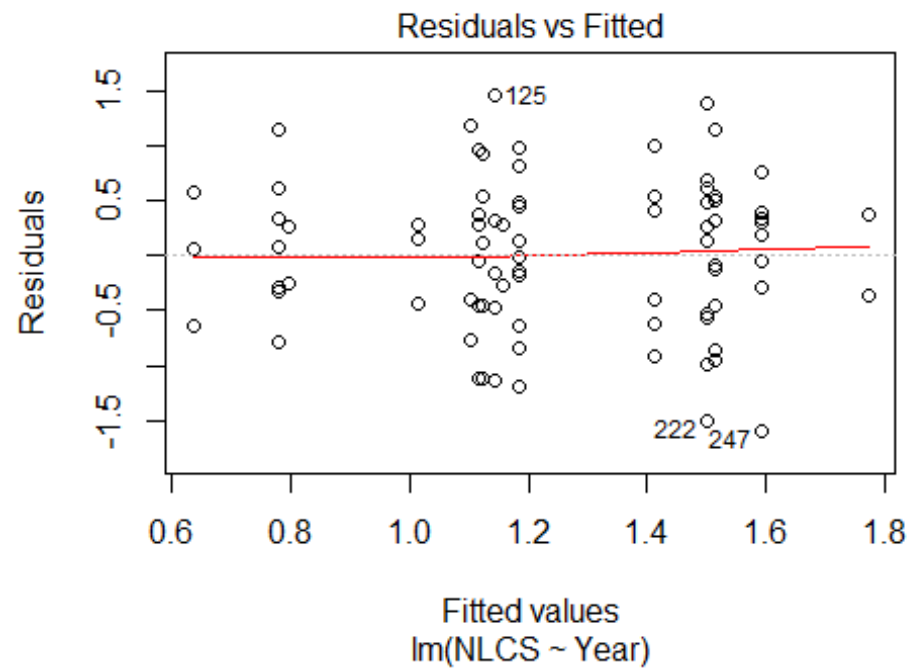
## Year2007          -0.52293      0.44929    -1.16   0.24557
## Year2008          -0.44206      0.44198    -1.00   0.31819
## Year2009          -0.63613      0.44954    -1.42   0.15829
## Year2010          -0.17038      0.43234    -0.39   0.69386
## Year2011          -0.50796      0.43770    -1.16   0.24694
## Year2012          -0.54011      0.43984    -1.23   0.22061
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.72
## Multiple R-squared:  0.0751, Adjusted R-squared:  0.0124
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 245 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.474  0.882  0.949  0.924  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      3.72e-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: 269"
## [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 2000 2001 2002 2003 2004 2005 2006 2007 2008 2011 2012
##    2    1    2    2    5    2    2    2    1    1    1    4    1    2
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2011 2012
##    1    0    0    1    2    1    1    0    1    1    1    1    0    1
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2011 2012
##    1    0    0    1    2    0    1    0    1    1    1    1    0    1

```

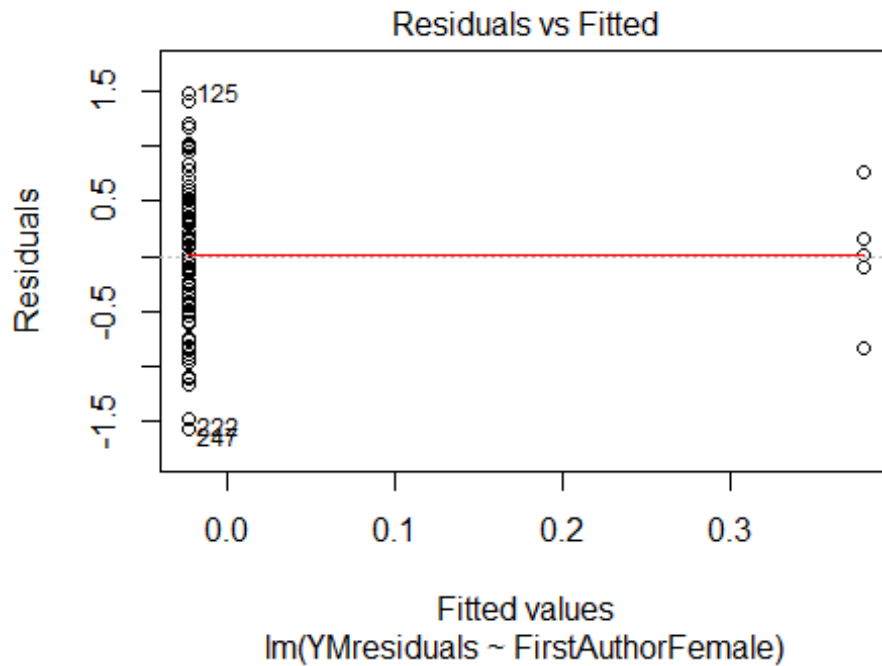
```

## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 10"
## [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
##    7    6    6    4    9    9    4   12    6   11    8   14   11   27   19
## 2011 2012
##   20   18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    0    2    2    5    3    0    8    2    6    3    5    6   12   10
## 2011 2012
##    9    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    0    1    1    4    3    0    4    2    6    2    4    4    6    7
## 2011 2012
##    7    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 4.6, df = 14, p-value = 1

```

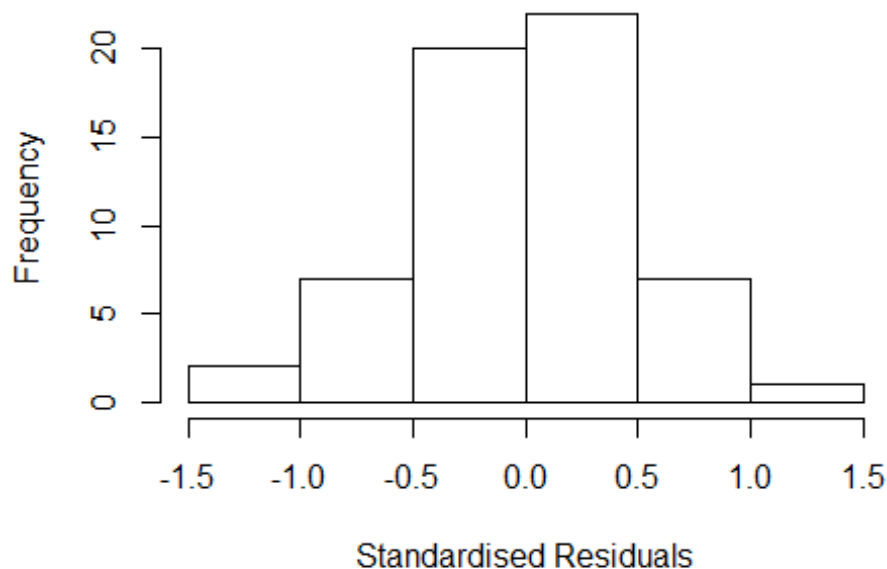


```
##  
## Bartlett test of homogeneity of variances  
##  
## data: YMresiduals by FirstAuthorFemale  
## Bartlett's K-squared = 0.16, df = 1, p-value = 0.7
```



```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 4.47213595499958"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 5.124e+00 1          2.264
## LastAuthorFemale  3.363e+00 1          1.834
## UniqueAuthors    9.374e+17 4         176.397
## Year              3.196e+18 14          4.580
```

Residuals from first and last author and team size



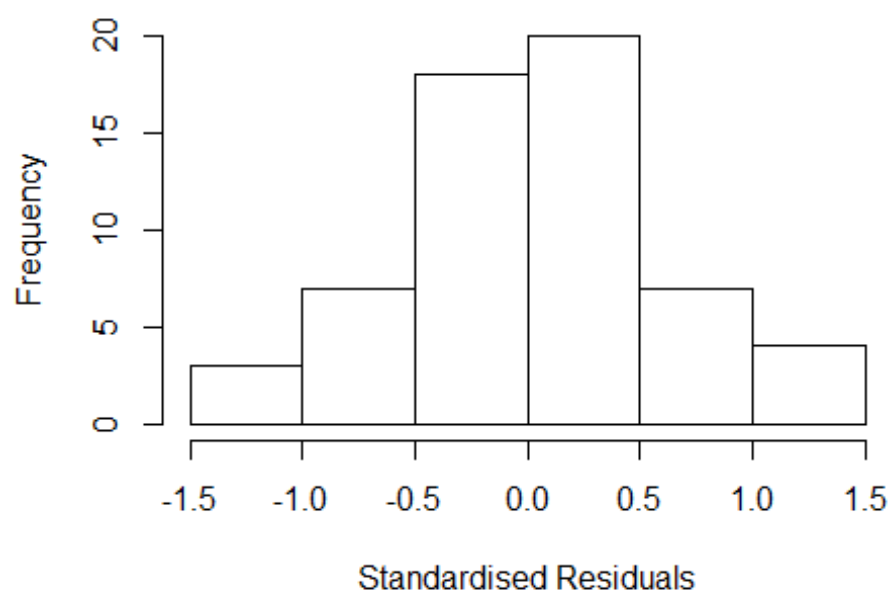
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3108 -0.2830  0.0325  0.2347  1.2644
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.77620    0.20533   3.78  0.00054 ***
## FirstAuthorFemale1 0.74278    0.21673   3.43  0.00148 **
## LastAuthorFemale1 -0.39658    0.25176  -1.58  0.12349
## UniqueAuthors2     0.32759    0.29269   1.12  0.27006
## UniqueAuthors3     0.26361    0.32024   0.82  0.41554
## UniqueAuthors4     1.68644    0.29630   5.69  1.5e-06 ***
## UniqueAuthors5    -0.18605    0.34319  -0.54  0.59090
## Year1998          0.62680    0.20533   3.05  0.00413 **
## Year1999          0.01918    0.29510   0.07  0.94851
## Year2000          0.50924    0.31188   1.63  0.11076
```

```

## Year2001      -0.45943    0.37914   -1.21  0.23307
## Year2003      -0.43819    0.47185   -0.93  0.35892
## Year2004       0.21700    0.16613    1.31  0.19933
## Year2005       0.20699    0.42542    0.49  0.62936
## Year2006      -0.52232    0.32555   -1.60  0.11690
## Year2007      -0.00733    0.44422   -0.02  0.98691
## Year2008       0.17211    0.35436    0.49  0.62998
## Year2009       0.33206    0.28878    1.15  0.25739
## Year2010       0.57304    0.26384    2.17  0.03617 *
## Year2011       0.82115    0.31790    2.58  0.01377 *
## Year2012       0.35555    0.35397    1.00  0.32150
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.575
## Multiple R-squared:  0.527, Adjusted R-squared:  0.278
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.582  0.894  0.969  0.925  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      1.69e-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 2.836 1 1.684
## LastAuthorFemale 8.145 1 2.854
## Year 12.297 14 1.094

```


Residuals from first and last author



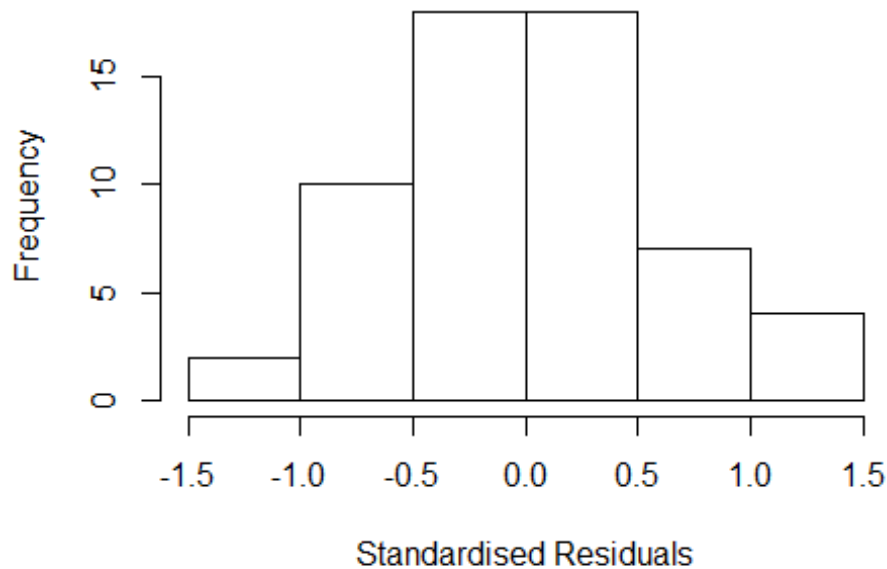
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2753 -0.3215 0.0139 0.3489 1.4366
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.940 0.270 3.49 0.0012 **
## FirstAuthorFemale1 0.777 0.245 3.18 0.0028 **
## LastAuthorFemale1 -0.419 0.195 -2.15 0.0377 *
## Year1998 0.463 0.270 1.72 0.0933 .
## Year1999 0.119 0.270 0.44 0.6612
## Year2000 0.592 0.404 1.47 0.1500
## Year2001 -0.299 0.409 -0.73 0.4693
## Year2003 -0.447 0.509 -0.88 0.3848
## Year2004 0.217 0.338 0.64 0.5243
## Year2005 0.120 0.414 0.29 0.7729
## Year2006 -0.422 0.302 -1.40 0.1694
## Year2007 0.295 0.640 0.46 0.6469
```

```

## Year2008          0.178      0.388      0.46      0.6481
## Year2009          0.307      0.348      0.88      0.3828
## Year2010          0.842      0.411      2.05      0.0468 *
## Year2011          0.833      0.300      2.78      0.0082 **
## Year2012          0.390      0.370      1.05      0.2990
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.711
## Multiple R-squared:  0.355, Adjusted R-squared:  0.109
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 50 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.663  0.914  0.967  0.932  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.69e-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 1.706 1          1.306
## Year              1.706 14          1.019

```

Residuals from first author



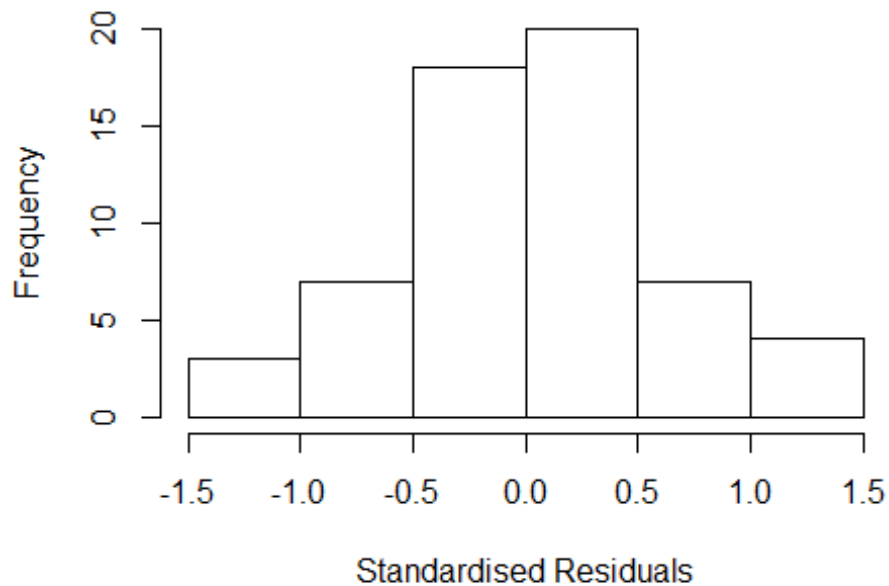
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.27e+00 -3.48e-01 2.36e-16 3.50e-01 1.43e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9400 0.2685 3.50 0.0011 **
## FirstAuthorFemale1 0.6900 0.2653 2.60 0.0127 *
## Year1998 0.4630 0.2685 1.72 0.0918 .
## Year1999 0.1190 0.2685 0.44 0.6598
## Year2000 0.4717 0.3800 1.24 0.2212
## Year2001 -0.2990 0.4067 -0.74 0.4661
## Year2003 -0.4354 0.5049 -0.86 0.3933
## Year2004 0.2170 0.3368 0.64 0.5228
## Year2005 0.0513 0.4082 0.13 0.9005
## Year2006 -0.4225 0.3011 -1.40 0.1678
## Year2007 0.2981 0.6283 0.47 0.6376
## Year2008 0.2032 0.3879 0.52 0.6030
```

```

## Year2009          0.3080      0.3464      0.89      0.3788
## Year2010          0.8396      0.4081      2.06      0.0457 *
## Year2011          0.7859      0.3018      2.60      0.0126 *
## Year2012          0.4032      0.3725      1.08      0.2850
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.742
## Multiple R-squared:  0.336, Adjusted R-squared:  0.105
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 52 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.692  0.921  0.972  0.939  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.69e-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 2.401 1          1.550
## Year            2.401 14          1.032

```

Residuals from last author



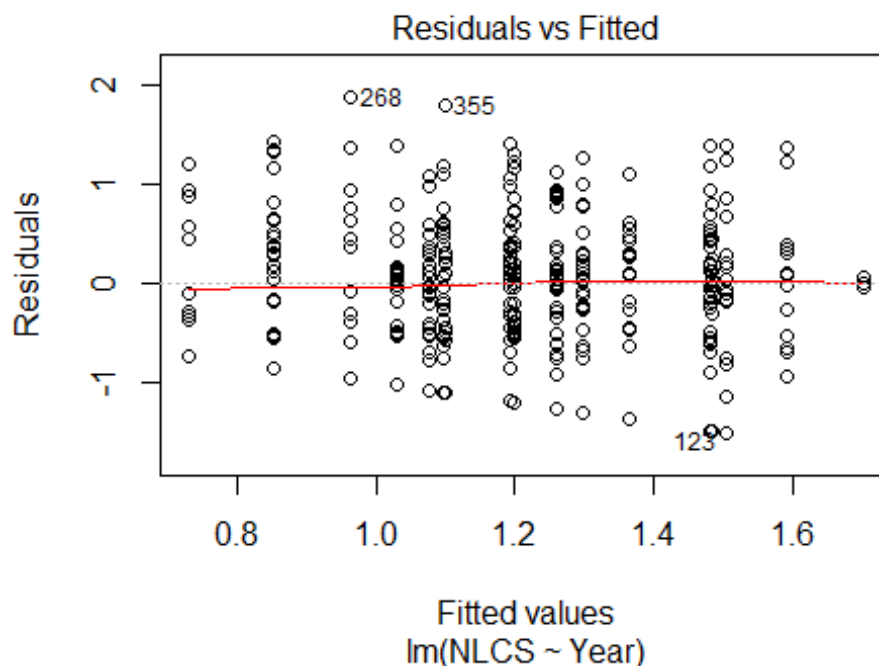
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.27233 -0.31826 -0.00726 0.35033 1.42430
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.940 0.268 3.50 0.0011 **
## LastAuthorFemale1 -0.204 0.231 -0.88 0.3813
## Year1998 0.463 0.268 1.73 0.0917 .
## Year1999 0.119 0.268 0.44 0.6597
## Year2000 0.529 0.402 1.32 0.1948
## Year2001 -0.299 0.406 -0.74 0.4658
## Year2003 -0.434 0.506 -0.86 0.3951
## Year2004 0.217 0.337 0.64 0.5226
## Year2005 0.231 0.395 0.59 0.5606
## Year2006 -0.423 0.301 -1.40 0.1677
## Year2007 0.298 0.627 0.48 0.6367
## Year2008 0.387 0.427 0.91 0.3703
```

```

## Year2009          0.308      0.346      0.89      0.3785
## Year2010          0.839      0.408      2.06      0.0457 *
## Year2011          0.917      0.299      3.06      0.0038 **
## Year2012          0.488      0.433      1.13      0.2654
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.746
## Multiple R-squared:  0.29,   Adjusted R-squared:  0.0424
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 49 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.695  0.899  0.961   0.931  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      1.69e-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: 59"
## [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
##   40   46   53    5   48   38   28   29   41   43   48   55   46   64   74
## 2011 2012
##   68   57
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   13   15   17    3   23   18    9   12   21   25   26   27   19   36   36
## 2011 2012
##   30   26
##

```

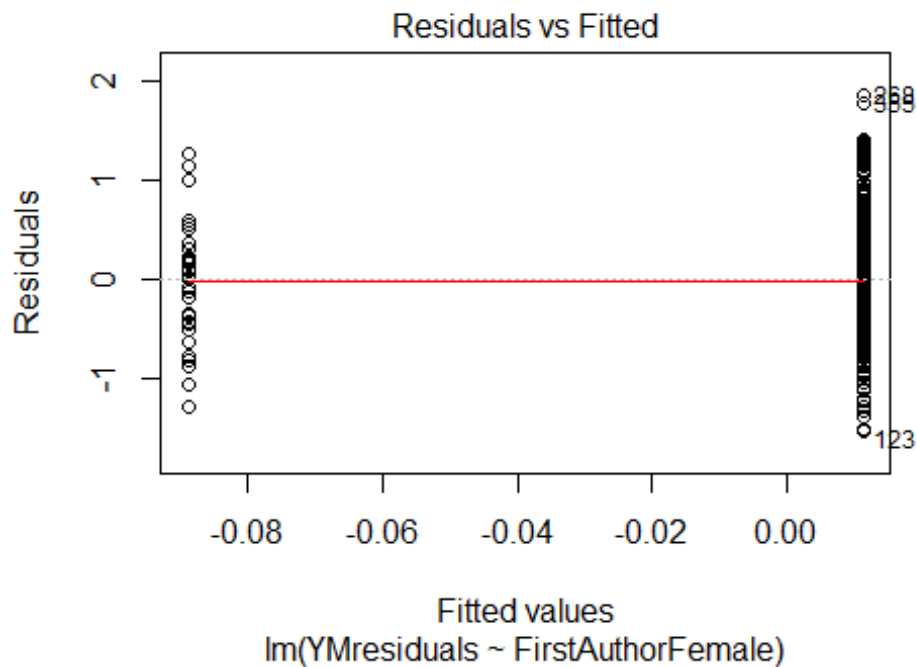
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    11    14    15     3    20    16     8     8    19    18    20    22    16    30    29
## 2011 2012
##    25    19
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 17, df = 16, p-value = 0.4
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.89, df = 1, p-value = 0.3
##
## [1] "Female first author team size 2018 geometric mean: 6"
## [1] "Male first author team size 2018 geometric mean: 2.20817902734762"
##
## 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 = 7, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.89897948556636"
## [1] "Male last author team size 2018 geometric mean: 2"

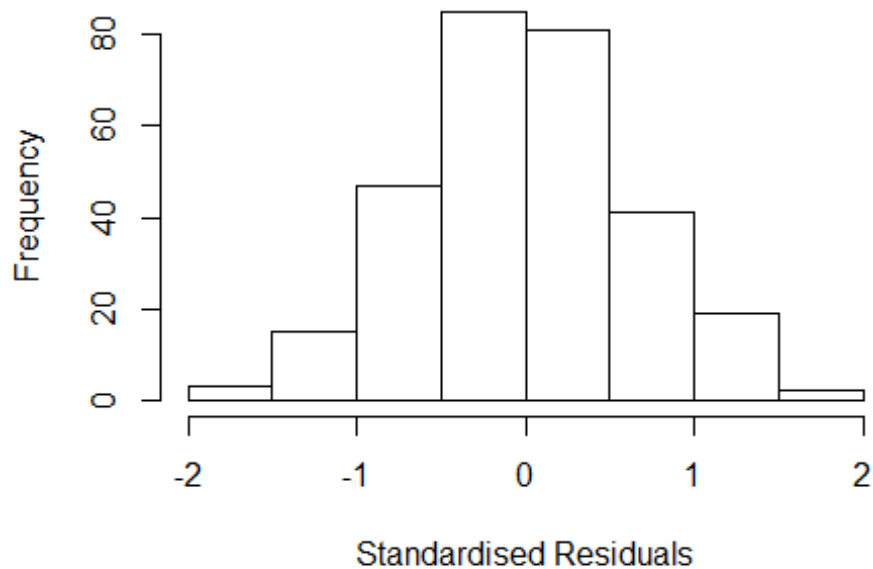
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 11, 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	1.478	1	1.216
LastAuthorFemale	1.588	1	1.260
UniqueAuthors	3.810	4	1.182
Year	6.328	16	1.059

Residuals from first and last author and team size



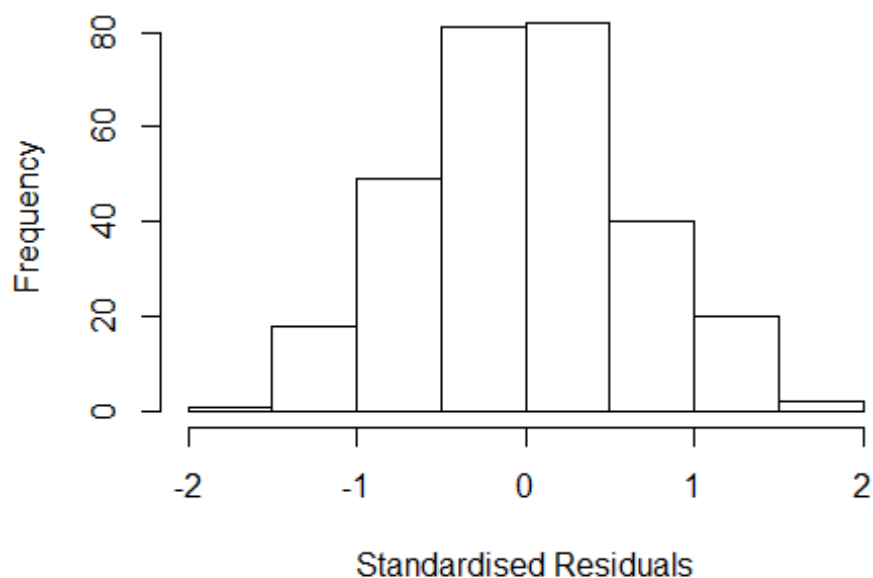
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.8390 -0.4327 -0.0265  0.4262  1.9813
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5953    0.2300    2.59  0.01017 *
## FirstAuthorFemale1  0.1645    0.1269    1.30  0.19587
## LastAuthorFemale1 -0.4055    0.1186   -3.42  0.00073 ***
## UniqueAuthors2     0.2029    0.1062    1.91  0.05708 .
## UniqueAuthors3     0.3646    0.1082    3.37  0.00086 ***
## UniqueAuthors4     0.3118    0.1727    1.81  0.07208 .
## UniqueAuthors5    -0.0431    0.1646   -0.26  0.79369
## Year1997          0.9241    0.2900    3.19  0.00161 **
## Year1998          0.8791    0.3055    2.88  0.00433 **
## Year1999          0.8764    0.2389    3.67  0.00029 ***
```

```

## Year2000          0.2499      0.2496      1.00  0.31769
## Year2001          0.2554      0.3760      0.68  0.49744
## Year2002          0.1727      0.3415      0.51  0.61347
## Year2003          0.7639      0.2831      2.70  0.00740 **
## Year2004          0.7567      0.2775      2.73  0.00682 **
## Year2005          0.5213      0.2695      1.93  0.05412 .
## Year2006          0.2046      0.2625      0.78  0.43630
## Year2007          0.3631      0.2626      1.38  0.16795
## Year2008          0.1411      0.2522      0.56  0.57633
## Year2009         -0.0116      0.2600     -0.04  0.96450
## Year2010          0.5760      0.2567      2.24  0.02567 *
## Year2011          0.5162      0.2668      1.93  0.05407 .
## Year2012          0.6784      0.2912      2.33  0.02055 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.635
## Multiple R-squared:  0.206, Adjusted R-squared:  0.141
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 275 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.310  0.879  0.953  0.908  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      3.41e-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 1.327 1      1.152
## LastAuthorFemale  1.416 1      1.190
## Year              1.658 16      1.016

```

Residuals from first and last author



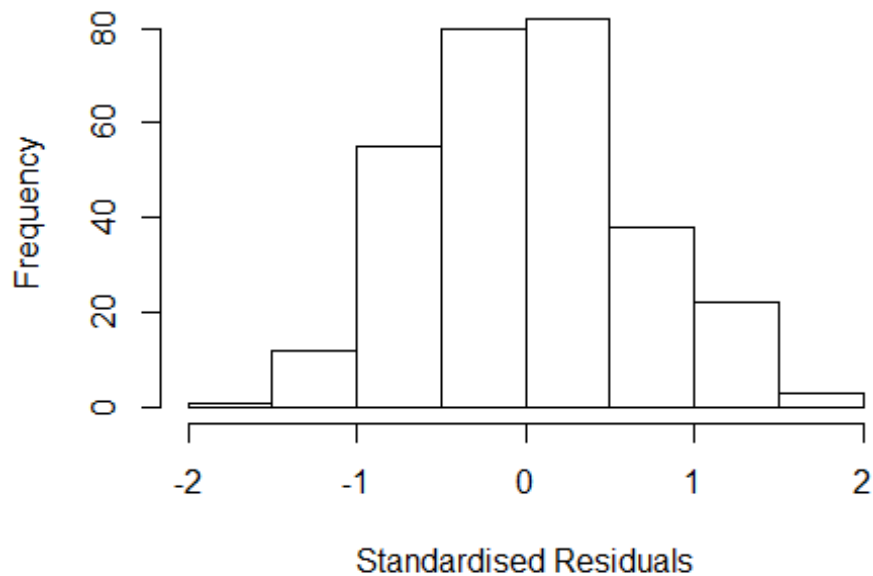
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.59846 -0.44890 -0.00785 0.43675 1.92475
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7163 0.2449 2.92 0.00374 **
## FirstAuthorFemale1 0.1678 0.1212 1.38 0.16745
## LastAuthorFemale1 -0.4490 0.1163 -3.86 0.00014 ***
## Year1997 0.9575 0.3071 3.12 0.00202 **
## Year1998 0.8822 0.3150 2.80 0.00546 **
## Year1999 0.9291 0.2546 3.65 0.00032 ***
## Year2000 0.3537 0.2737 1.29 0.19740
## Year2001 0.2965 0.3517 0.84 0.39993
## Year2002 0.2440 0.3559 0.69 0.49357
## Year2003 0.7721 0.3048 2.53 0.01187 *
## Year2004 0.7567 0.2828 2.68 0.00790 **
## Year2005 0.5322 0.2780 1.91 0.05658 .
```

```

## Year2006          0.3186      0.2910      1.09  0.27460
## Year2007          0.4238      0.2886      1.47  0.14305
## Year2008          0.2247      0.2666      0.84  0.40014
## Year2009          0.0726      0.2796      0.26  0.79524
## Year2010          0.6039      0.2786      2.17  0.03104 *
## Year2011          0.5955      0.2841      2.10  0.03700 *
## Year2012          0.7775      0.3039      2.56  0.01104 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.667
## Multiple R-squared:  0.16,   Adjusted R-squared:  0.105
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 263 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.385  0.875   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      3.41e-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 1.459 1      1.208
## Year              1.459 16      1.012

```

Residuals from first author



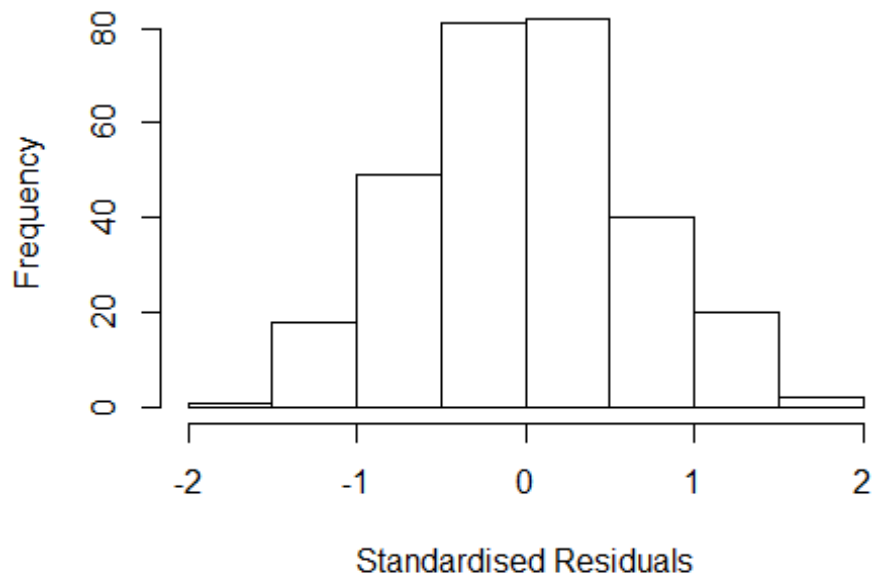
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5986 -0.4369 -0.0181 0.4409 1.9262
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7158 0.2454 2.92 0.00383 **
## FirstAuthorFemale1 0.0236 0.1250 0.19 0.85052
## Year1997 0.9165 0.3038 3.02 0.00279 **
## Year1998 0.8828 0.3154 2.80 0.00548 **
## Year1999 0.9777 0.2508 3.90 0.00012 ***
## Year2000 0.3542 0.2742 1.29 0.19750
## Year2001 0.2454 0.3722 0.66 0.51028
## Year2002 0.2431 0.3559 0.68 0.49516
## Year2003 0.7733 0.3050 2.54 0.01178 *
## Year2004 0.7209 0.2775 2.60 0.00989 **
## Year2005 0.5119 0.2875 1.78 0.07609 .
## Year2006 0.2718 0.2937 0.93 0.35549
```

```

## Year2007          0.4454      0.2906      1.53  0.12653
## Year2008          0.2238      0.2688      0.83  0.40595
## Year2009          0.0414      0.2799      0.15  0.88251
## Year2010          0.5989      0.2802      2.14  0.03343 *
## Year2011          0.5426      0.2844      1.91  0.05746 .
## Year2012          0.7203      0.3038      2.37  0.01843 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.663
## Multiple R-squared:  0.139, Adjusted R-squared:  0.0858
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 265 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.379  0.867   0.952   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      3.41e-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 1.388 1          1.178
## Year              1.388 16          1.010

```

Residuals from last author



```
## [1] "List of 0 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
## -1.59859 -0.43694 -0.00733 0.42254 1.92608
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7158 0.2454 2.92 0.00382 **
## LastAuthorFemale1 -0.3802 0.1130 -3.36 0.00088 ***
## Year1997 0.9516 0.3066 3.10 0.00211 **
## Year1998 0.8828 0.3153 2.80 0.00548 **
## Year1999 0.9855 0.2468 3.99 8.4e-05 ***
## Year2000 0.3542 0.2741 1.29 0.19748
## Year2001 0.3078 0.3560 0.86 0.38802
## Year2002 0.2431 0.3559 0.68 0.49508
## Year2003 0.7733 0.3050 2.54 0.01178 *
## Year2004 0.7902 0.2815 2.81 0.00536 **
## Year2005 0.5479 0.2780 1.97 0.04978 *
## Year2006 0.3224 0.2947 1.09 0.27484
```

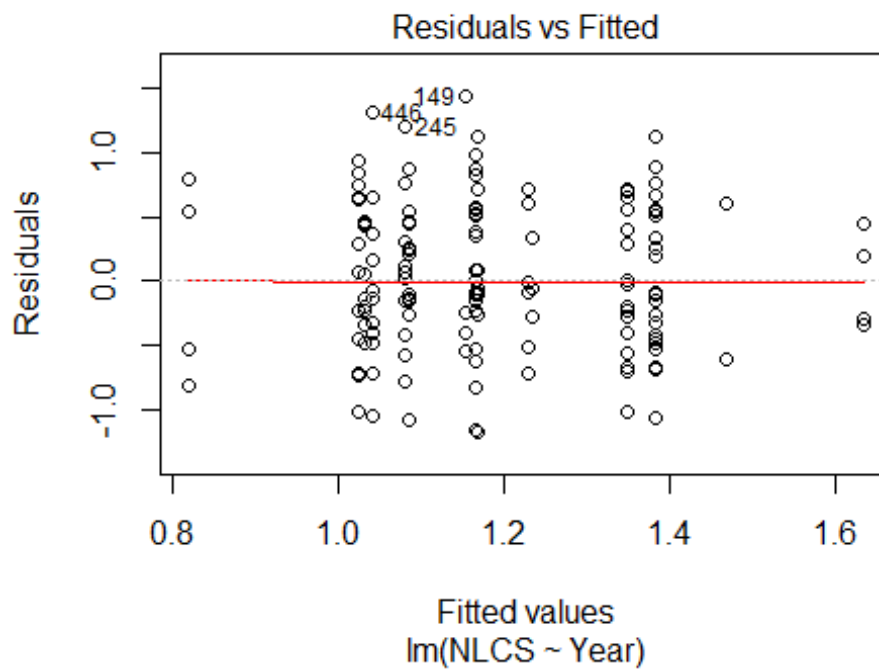
```

## Year2007          0.4487      0.2901      1.55  0.12307
## Year2008          0.2537      0.2651      0.96  0.33949
## Year2009          0.0907      0.2785      0.33  0.74498
## Year2010          0.6165      0.2782      2.22  0.02750 *
## Year2011          0.6027      0.2839      2.12  0.03465 *
## Year2012          0.7983      0.3044      2.62  0.00920 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.663
## Multiple R-squared:  0.157, Adjusted R-squared:  0.105
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 265 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.379  0.870   0.953   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      3.41e-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: 293"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##   15   11   23   15   39   24   13   13   18   30   26   34   23   34   28
## 2012
##   26
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    6    3    9    4   14    5    2    4    9   12   13   25   10   21   12
## 2012

```



```
##      7
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    5    2    8    4   13    4    1    3    6    8   12   24   10   15    9
## 2012
##    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 8.4, df = 15, p-value = 0.9
```



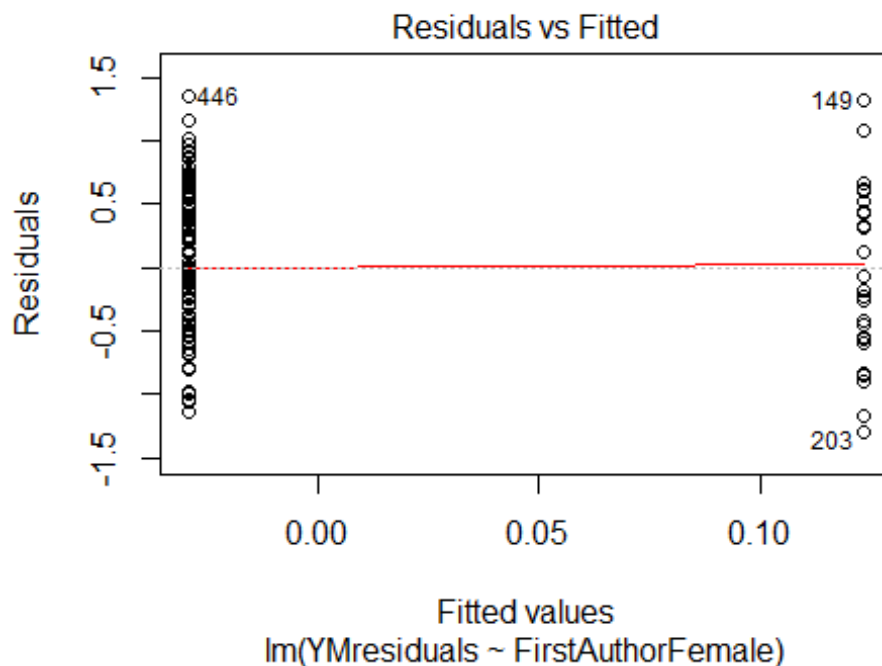
```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.3, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 5.09164336965949"
## [1] "Male first author team size 2018 geometric mean: 2.10109212965504"

## 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 = 35, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.28942848510666"
## [1] "Male last author team size 2018 geometric mean: 2.52669153631783"

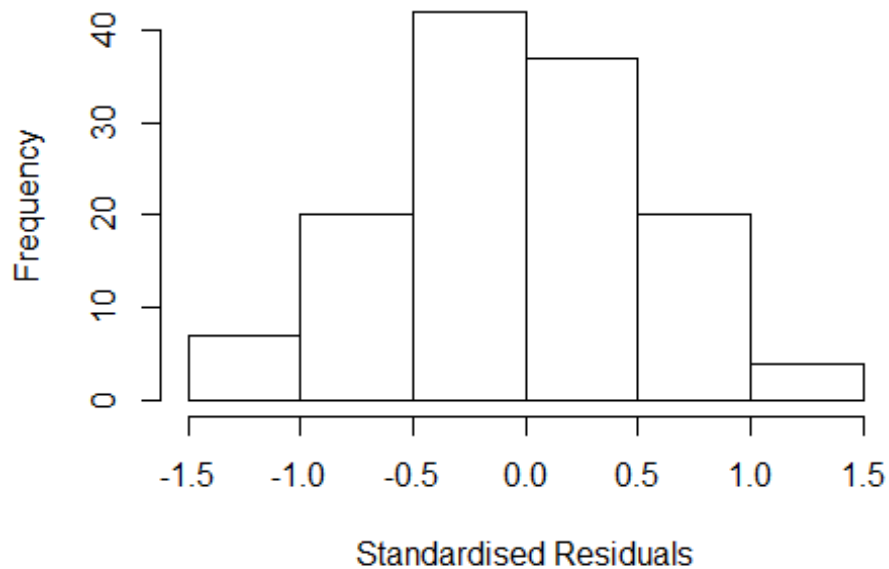
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16, 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	3.668	1	1.915
LastAuthorFemale	2.491	1	1.578
UniqueAuthors	29.279	4	1.525
Year	84.631	15	1.159

Residuals from first and last author and team size



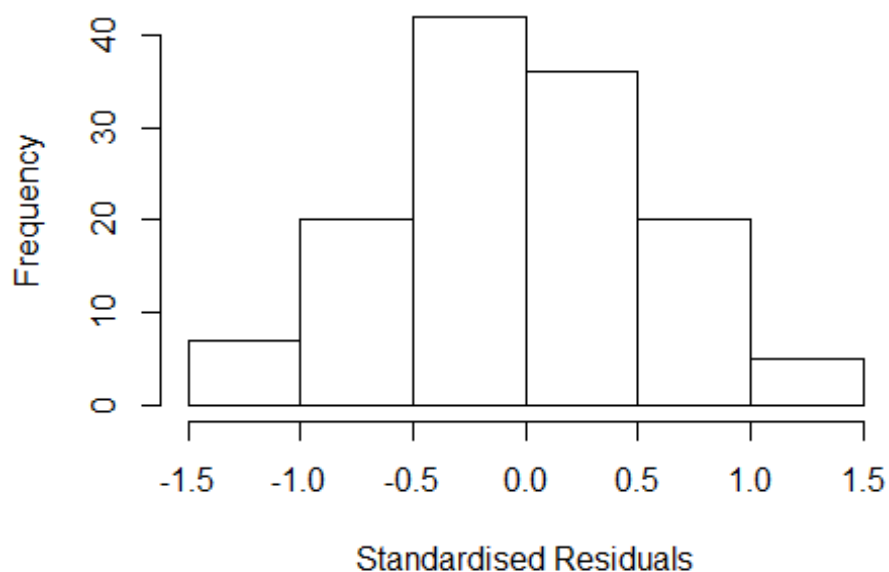
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2524 -0.2837 -0.0413 0.3662 1.2842
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0951 0.2080 5.26 7.2e-07 ***
## FirstAuthorFemale1 0.2946 0.1949 1.51 0.134
## LastAuthorFemale1 -0.2336 0.1446 -1.62 0.109
## UniqueAuthors2 0.1176 0.1638 0.72 0.474
## UniqueAuthors3 0.1431 0.2125 0.67 0.502
## UniqueAuthors4 -0.5194 0.2245 -2.31 0.023 *
## UniqueAuthors5 -0.0395 0.2530 -0.16 0.876
## Year1998 0.2789 0.2496 1.12 0.266
## Year1999 -0.0396 0.2714 -0.15 0.884
## Year2000 -0.2876 0.4971 -0.58 0.564
```

```

## Year2001          0.0428      0.2563      0.17      0.868
## Year2002         -0.0645      0.3896     -0.17      0.869
## Year2003          0.9759      0.2080      4.69    8.0e-06 ***
## Year2004          0.2663      0.2674      1.00      0.322
## Year2005         -0.0910      0.2799     -0.33      0.746
## Year2006         -0.0697      0.2985     -0.23      0.816
## Year2007          0.0145      0.3241      0.04      0.964
## Year2008          0.2882      0.2657      1.08      0.281
## Year2009          0.1377      0.3028      0.45      0.650
## Year2010          0.1321      0.3136      0.42      0.674
## Year2011         -0.0213      0.3802     -0.06      0.955
## Year2012          0.1738      0.3181      0.55      0.586
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.556
## Multiple R-squared:  0.198, Adjusted R-squared:  0.0417
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.574  0.846  0.955  0.907  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      7.69e-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 4.343  1      2.084
## LastAuthorFemale  2.572  1      1.604
## Year              6.486 15      1.064

```

Residuals from first and last author



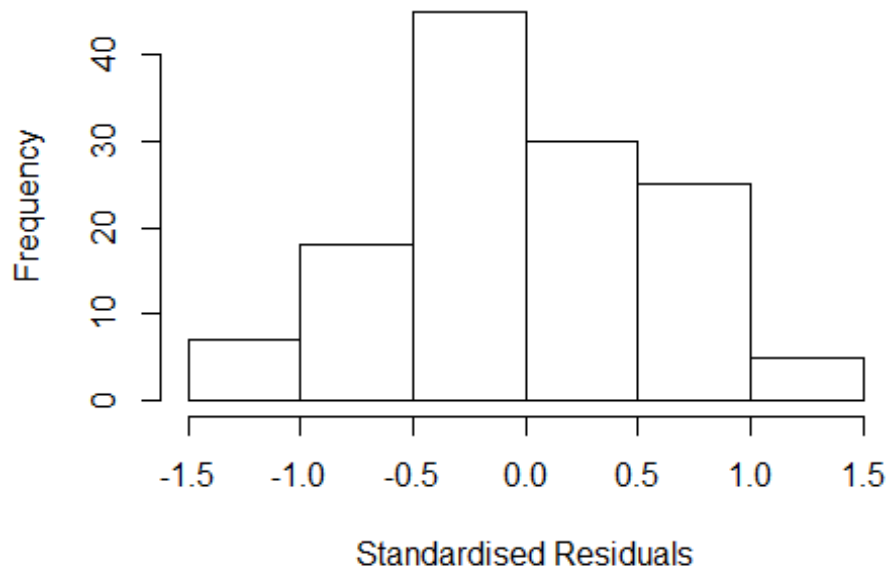
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3595 -0.3516 -0.0458 0.4285 1.2890
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1377 0.2229 5.10 1.4e-06 ***
## FirstAuthorFemale1 0.2850 0.2127 1.34 0.18
## LastAuthorFemale1 -0.2173 0.1621 -1.34 0.18
## Year1998 0.2363 0.2620 0.90 0.37
## Year1999 -0.1447 0.2693 -0.54 0.59
## Year2000 -0.3321 0.4907 -0.68 0.50
## Year2001 -0.0318 0.2792 -0.11 0.91
## Year2002 -0.0638 0.4049 -0.16 0.88
## Year2003 0.9333 0.2229 4.19 5.6e-05 ***
## Year2004 0.3103 0.2901 1.07 0.29
## Year2005 -0.0631 0.3322 -0.19 0.85
## Year2006 -0.0583 0.3173 -0.18 0.85
```

```

## Year2007          -0.2661      0.3342    -0.80      0.43
## Year2008           0.2251      0.2657      0.85      0.40
## Year2009           0.1842      0.2912      0.63      0.53
## Year2010           0.1353      0.2914      0.46      0.64
## Year2011          -0.0686      0.3619     -0.19      0.85
## Year2012           0.2141      0.2827      0.76      0.45
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.578
## Multiple R-squared:  0.12,   Adjusted R-squared:  -0.0141
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 117 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.559  0.857   0.945   0.904   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      7.69e-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 2.89 1          1.700
## Year              2.89 15          1.036

```

Residuals from first author



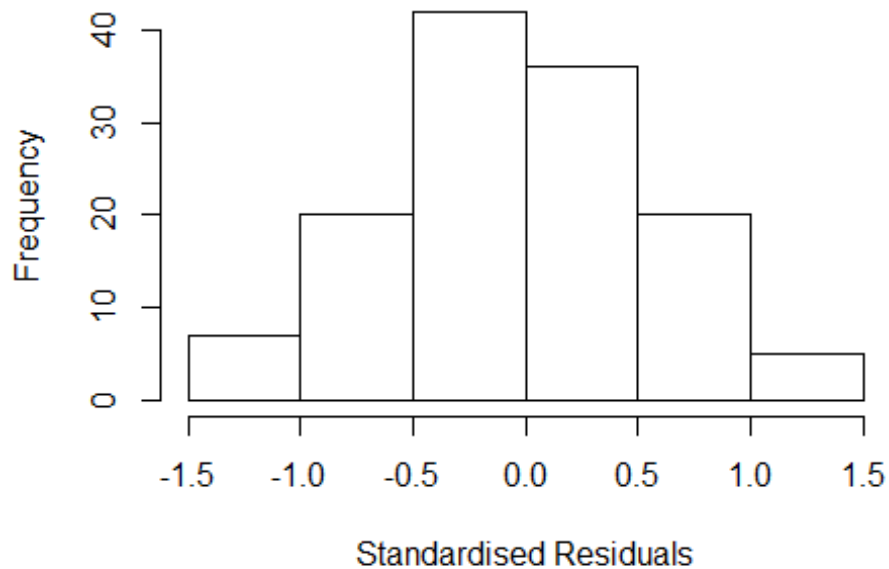
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2462 -0.3979 -0.0787 0.4491 1.3074
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09225 0.23284 4.69 7.7e-06 ***
## FirstAuthorFemale1 0.19562 0.18874 1.04 0.30
## Year1998 0.28175 0.27032 1.04 0.30
## Year1999 -0.13651 0.27886 -0.49 0.63
## Year2000 -0.31814 0.45082 -0.71 0.48
## Year2001 -0.00719 0.28865 -0.02 0.98
## Year2002 0.00782 0.42657 0.02 0.99
## Year2003 0.97875 0.23284 4.20 5.3e-05 ***
## Year2004 0.34103 0.32024 1.06 0.29
## Year2005 -0.07924 0.34147 -0.23 0.82
## Year2006 -0.06258 0.32976 -0.19 0.85
## Year2007 -0.20520 0.33987 -0.60 0.55
```

```

## Year2008          0.24925    0.26797    0.93    0.35
## Year2009          0.21450    0.29979    0.72    0.48
## Year2010          0.12768    0.30217    0.42    0.67
## Year2011         -0.04165    0.35044   -0.12    0.91
## Year2012          0.14930    0.32316    0.46    0.64
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.636
## Multiple R-squared:  0.0987, Adjusted R-squared:  -0.0289
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 123 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.653  0.890  0.953  0.923  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      7.69e-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 1.869 1          1.367
## Year            1.869 15          1.021

```


Residuals from last author



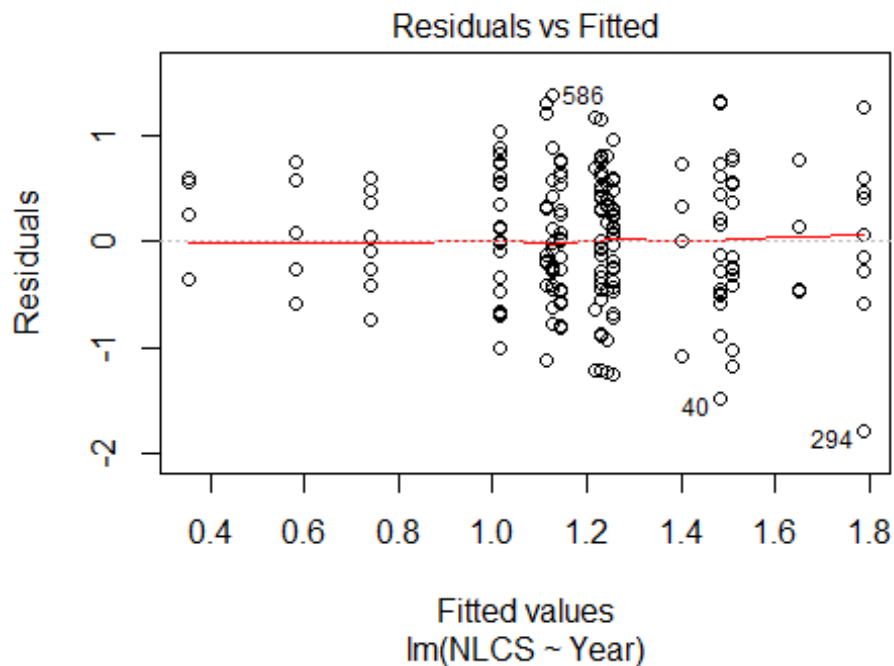
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2467 -0.3617 -0.0602 0.4682 1.5158
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1174 0.2266 4.93 2.8e-06 ***
## LastAuthorFemale1 -0.1208 0.1389 -0.87 0.39
## Year1998 0.2566 0.2651 0.97 0.34
## Year1999 -0.0146 0.2638 -0.06 0.96
## Year2000 -0.2741 0.5271 -0.52 0.60
## Year2001 0.0449 0.2758 0.16 0.87
## Year2002 -0.0412 0.4448 -0.09 0.93
## Year2003 0.9536 0.2266 4.21 5.2e-05 ***
## Year2004 0.4853 0.3031 1.60 0.11
## Year2005 -0.0670 0.3186 -0.21 0.83
## Year2006 -0.0246 0.3234 -0.08 0.94
## Year2007 -0.1566 0.3299 -0.47 0.64
```

```

## Year2008          0.3060      0.2537      1.21      0.23
## Year2009          0.2609      0.2987      0.87      0.38
## Year2010          0.1293      0.2942      0.44      0.66
## Year2011         -0.0439      0.3412     -0.13      0.90
## Year2012          0.1856      0.2997      0.62      0.54
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.0948, Adjusted R-squared:  -0.0333
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 119 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.517  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      7.69e-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: 130"
## [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
##   33   42   32   22   26   45   38   33   32   38   32   33   66   57   47
## 2011 2012
##   42   55
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   15    4    9    6    4    9   10    9   13    8    7   15   17   28
## 2011 2012
##   20   29

```

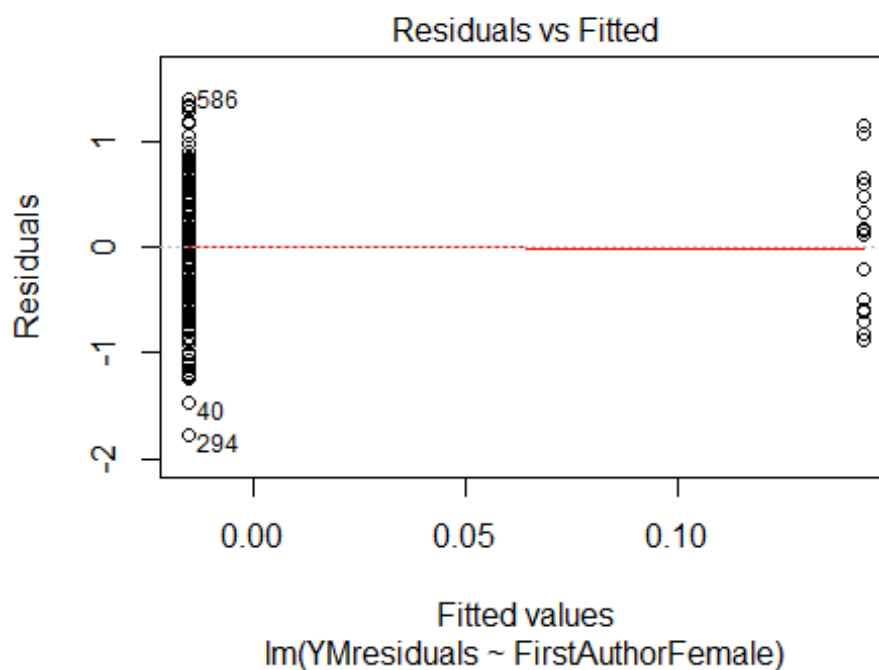
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   13    4    8    5    4    8    8    8   12    7    7   12   15   26
## 2011 2012
##   17   24
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 17, df = 16, p-value = 0.4
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.042, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 3.7224194364084"
## [1] "Male first author team size 2018 geometric mean: 2.80844242243156"
## 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 = 52, p-value = 0.3
## 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: 3.06064971804522"

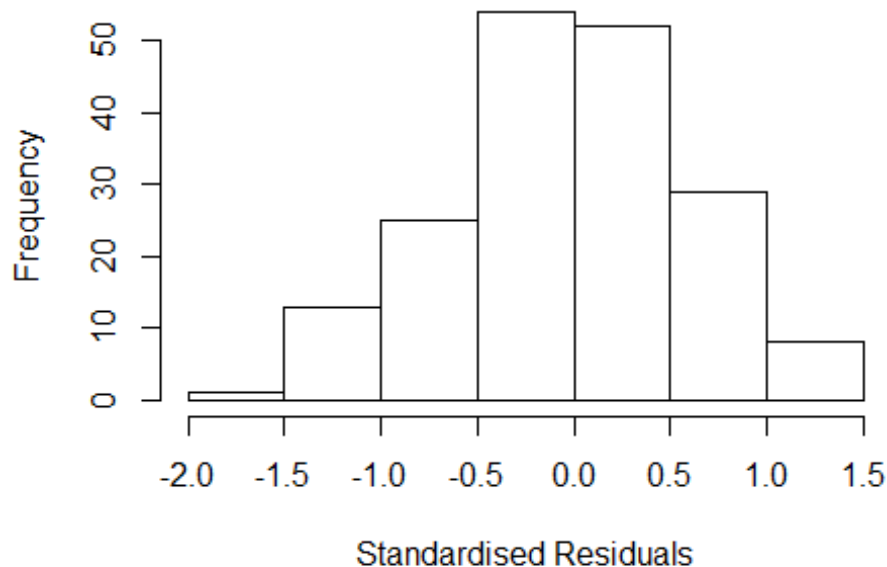
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7, 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	2.359	1	1.536
LastAuthorFemale	2.175	1	1.475
UniqueAuthors	3.920	4	1.186
Year	4.503	16	1.048

Residuals from first and last author and team size



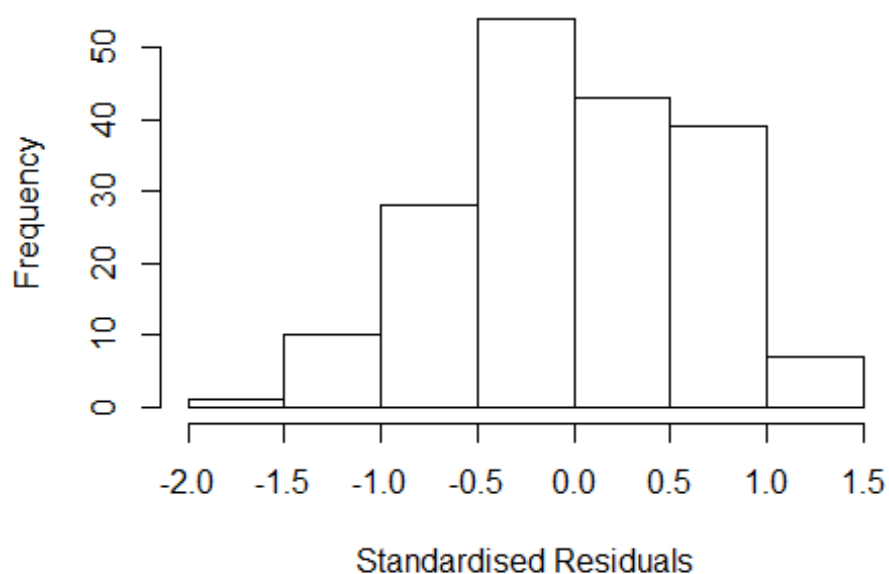
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.7372 -0.3969 -0.0261 0.4211 1.3517
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2234 0.7738 1.58 0.116
## FirstAuthorFemale1 0.2824 0.1900 1.49 0.139
## LastAuthorFemale1 -0.4848 0.1996 -2.43 0.016 *
## UniqueAuthors2 0.0685 0.1373 0.50 0.619
## UniqueAuthors3 0.3303 0.1676 1.97 0.051 .
## UniqueAuthors4 0.0213 0.2152 0.10 0.921
## UniqueAuthors5 0.3601 0.2134 1.69 0.094 .
## Year1997 0.2229 0.8304 0.27 0.789
## Year1998 0.4354 0.8209 0.53 0.597
## Year1999 -0.0186 0.8528 -0.02 0.983
```

```

## Year2000          -0.8296      0.7966    -1.04      0.299
## Year2001           0.1886      0.8525      0.22      0.825
## Year2002           0.4454      0.8243      0.54      0.590
## Year2003          -0.1382      0.8469     -0.16      0.871
## Year2004          -0.1770      0.8286     -0.21      0.831
## Year2005           0.1941      0.8104      0.24      0.811
## Year2006          -0.4734      0.7987     -0.59      0.554
## Year2007          -0.9805      0.8058     -1.22      0.225
## Year2008          -0.2682      0.8090     -0.33      0.741
## Year2009          -0.3014      0.8022     -0.38      0.708
## Year2010          -0.0894      0.7946     -0.11      0.911
## Year2011          -0.2256      0.7992     -0.28      0.778
## Year2012          -0.1022      0.7890     -0.13      0.897
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.646
## Multiple R-squared:  0.224, Adjusted R-squared:  0.116
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 172 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.450  0.900  0.954  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      5.49e-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 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.640 1      1.281
## LastAuthorFemale  1.809 1      1.345
## Year              1.411 16      1.011

```

Residuals from first and last author



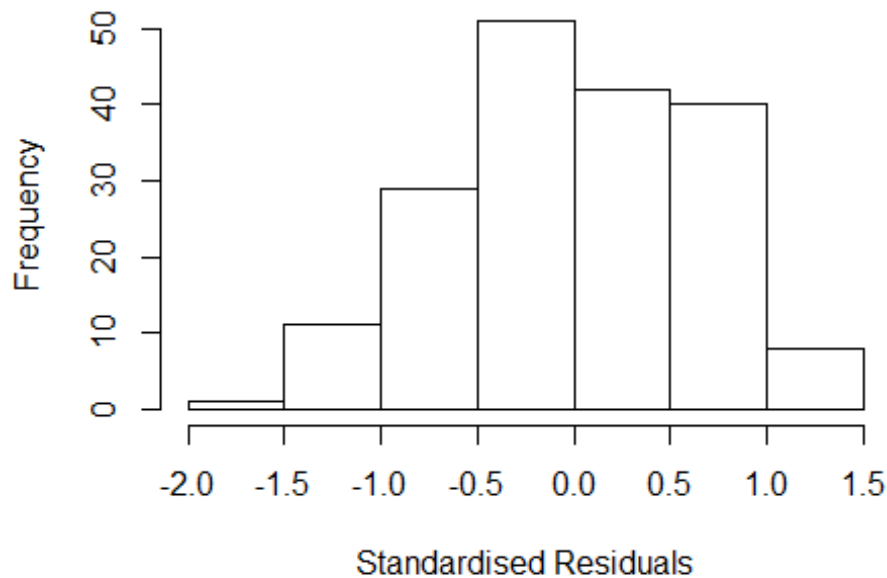
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.8049 -0.4455 -0.0216 0.5033 1.4620
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.23052 0.75037 1.64 0.103
## FirstAuthorFemale1 0.26212 0.17449 1.50 0.135
## LastAuthorFemale1 -0.44536 0.21057 -2.12 0.036 *
## Year1997 0.26343 0.79038 0.33 0.739
## Year1998 0.45442 0.79500 0.57 0.568
## Year1999 -0.01635 0.82869 -0.02 0.984
## Year2000 -0.78504 0.78672 -1.00 0.320
## Year2001 0.27773 0.84689 0.33 0.743
## Year2002 0.57439 0.80361 0.71 0.476
## Year2003 -0.00374 0.82445 0.00 0.996
## Year2004 -0.02604 0.79138 -0.03 0.974
## Year2005 0.22914 0.77402 0.30 0.768
```

```

## Year2006      -0.42354    0.76764   -0.55    0.582
## Year2007      -0.92217    0.77336   -1.19    0.235
## Year2008      -0.18153    0.77194   -0.24    0.814
## Year2009      -0.16827    0.77276   -0.22    0.828
## Year2010       0.03563    0.76112    0.05    0.963
## Year2011      -0.10776    0.76182   -0.14    0.888
## Year2012       0.00747    0.75575    0.01    0.992
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.685
## Multiple R-squared:  0.195, Adjusted R-squared:  0.106
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 167 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.467  0.901  0.950  0.922  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      5.49e-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"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.16 1      1.077
## Year              1.16 16      1.005

```


Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.802 -0.448 -0.017 0.540 1.460
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22994 0.73556 1.67 0.096 .
## FirstAuthorFemale1 0.11579 0.17986 0.64 0.521
## Year1997 0.23112 0.77933 0.30 0.767
## Year1998 0.37673 0.79275 0.48 0.635
## Year1999 -0.06924 0.79959 -0.09 0.931
## Year2000 -0.78380 0.77137 -1.02 0.311
## Year2001 0.19252 0.81534 0.24 0.814
## Year2002 0.57233 0.79432 0.72 0.472
## Year2003 -0.05053 0.80989 -0.06 0.950
## Year2004 -0.00883 0.78294 -0.01 0.991
## Year2005 0.22916 0.76052 0.30 0.764
## Year2006 -0.45612 0.75994 -0.60 0.549
```

```

## Year2007          -0.90125    0.75633   -1.19    0.235
## Year2008          -0.17924    0.75459   -0.24    0.813
## Year2009          -0.17749    0.75487   -0.24    0.814
## Year2010          -0.01737    0.74854   -0.02    0.982
## Year2011          -0.08911    0.74778   -0.12    0.905
## Year2012          -0.05308    0.74312   -0.07    0.943
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.695
## Multiple R-squared:  0.168, Adjusted R-squared:  0.0813
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.482  0.901  0.946  0.922  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      5.49e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.268 1          1.126
## Year              1.268 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
## -1.8024 -0.4429 -0.0404  0.5176  1.4604

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.22997    0.73708    1.67  0.097 .
## LastAuthorFemale1 -0.33852    0.17004   -1.99  0.048 *
## Year1997        0.27818    0.78058    0.36  0.722
## Year1998        0.49797    0.77785    0.64  0.523
## Year1999       -0.02703    0.81136   -0.03  0.973
## Year2000       -0.78385    0.77381   -1.01  0.313
## Year2001        0.32477    0.84429    0.38  0.701
## Year2002        0.57242    0.79172    0.72  0.471
## Year2003        0.02725    0.83506    0.03  0.974
## Year2004        0.00429    0.78460    0.01  0.996
## Year2005        0.22916    0.76112    0.30  0.764
## Year2006       -0.40291    0.75249   -0.54  0.593
## Year2007       -0.88449    0.75629   -1.17  0.244
## Year2008       -0.17934    0.75869   -0.24  0.813
## Year2009       -0.15252    0.75719   -0.20  0.841
## Year2010        0.05034    0.74758    0.07  0.946
## Year2011       -0.07550    0.74942   -0.10  0.920
## Year2012        0.03147    0.74256    0.04  0.966
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.695
## Multiple R-squared:  0.184, Adjusted R-squared:  0.0994
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.481  0.903   0.951   0.923   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      5.49e-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: 182"
## [1] ""

```

```

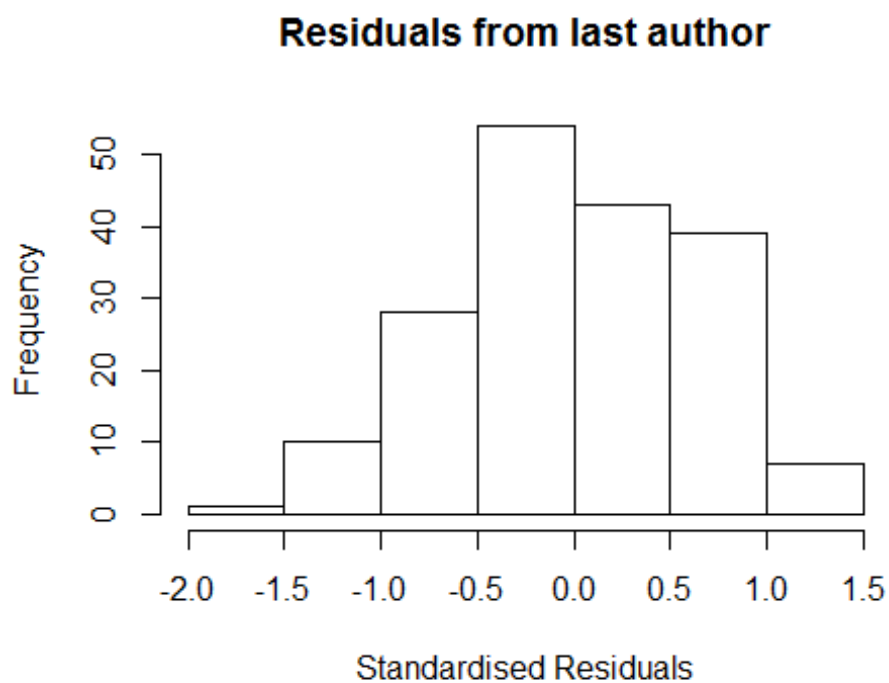
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2211"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 26 15 7 10 27 27 19 21 15 14 29 40 39 30
## 2011 2012
## 37 54
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 7 3 1 2 4 6 6 4 5 2 7 7 12 15
## 2011 2012
## 12 29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 6 3 1 1 4 3 4 3 5 2 7 6 10 13
## 2011 2012
## 11 24
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.09534502215844"
## [1] "Male first author team size 2018 geometric mean: 3.34683583453297"

## 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 = 53, p-value = 0.4
## 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: 3.70602951009135"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

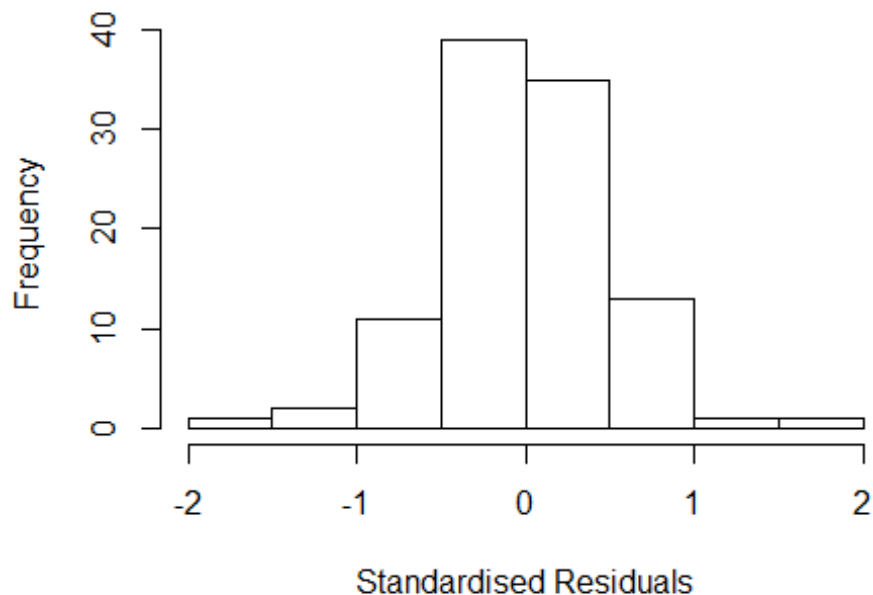
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, 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"
##
```

	GVI	F	Df	$GVI^{\frac{1}{2 \cdot Df}}$
FirstAuthorFemale	7.841e+00	1		2.800
LastAuthorFemale	6.979e+00	1		2.642
UniqueAuthors	4.319e+15	4		90.037
Year	2.122e+16	15		3.501

Residuals from first and last author and team size



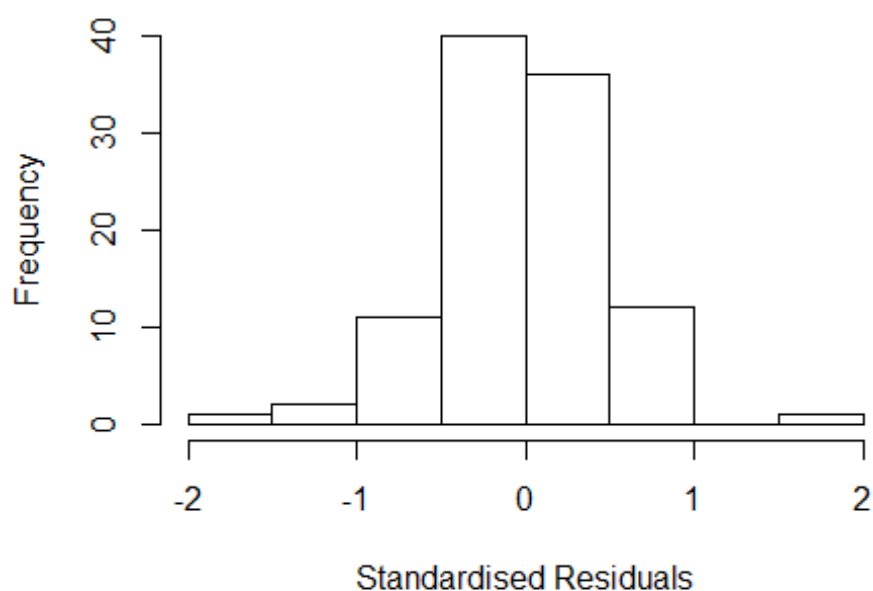
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.73e+00 -2.84e-01 -8.05e-16 3.04e-01 1.88e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.32439 0.25427 5.21 1.4e-06 ***
## FirstAuthorFemale1 0.18299 0.25018 0.73 0.4666
## LastAuthorFemale1 -0.35306 0.29485 -1.20 0.2346
## UniqueAuthors2 0.07248 0.14128 0.51 0.6093
## UniqueAuthors3 -0.03807 0.25780 -0.15 0.8830
## UniqueAuthors4 0.09513 0.29083 0.33 0.7444
## UniqueAuthors5 0.12189 0.19258 0.63 0.5286
## Year1998 0.14046 0.31996 0.44 0.6618
## Year1999 0.50612 0.19787 2.56 0.0124 *
## Year2000 -0.15239 0.25427 -0.60 0.5506
```

```

## Year2001      -0.24588    0.43206   -0.57    0.5709
## Year2002      0.58835    0.33009    1.78    0.0784 .
## Year2003      0.33492    1.04097    0.32    0.7485
## Year2004      0.20479    0.24133    0.85    0.3986
## Year2005     -0.00419    0.32938   -0.01    0.9899
## Year2006     -0.53635    0.37758   -1.42    0.1593
## Year2007     -0.93793    0.27703   -3.39    0.0011 **
## Year2008     -0.65054    0.23089   -2.82    0.0061 **
## Year2009      0.02429    0.30700    0.08    0.9371
## Year2010      0.05782    0.23929    0.24    0.8097
## Year2011     -0.17779    0.26612   -0.67    0.5060
## Year2012     -0.14355    0.22370   -0.64    0.5229
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.364, Adjusted R-squared:  0.2
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 92 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0893 0.8910 0.9580 0.9020 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          9.71e-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 -1.128e+01 1      NaN
## LastAuthorFemale  1.108e-13 1      3.328e-07
## Year              3.283e-13 15      3.836e-01

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.6525 -0.2742 -0.0058 0.3196 1.7474
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3952 0.1938 7.20 2.3e-10 ***
## FirstAuthorFemale1 0.2152 0.2279 0.94 0.34762
## LastAuthorFemale1 -0.3673 0.2685 -1.37 0.17489
## Year1998 0.0760 0.2674 0.28 0.77691
## Year1999 0.5078 0.1938 2.62 0.01042 *
## Year2000 -0.2232 0.1938 -1.15 0.25267
## Year2001 -0.2623 0.4171 -0.63 0.53110
## Year2002 0.5070 0.2830 1.79 0.07677 .
## Year2003 0.2572 1.0516 0.24 0.80734
## Year2004 0.2120 0.2358 0.90 0.37119
## Year2005 -0.0244 0.3388 -0.07 0.94267
## Year2006 -0.6423 0.3649 -1.76 0.08198 .
```

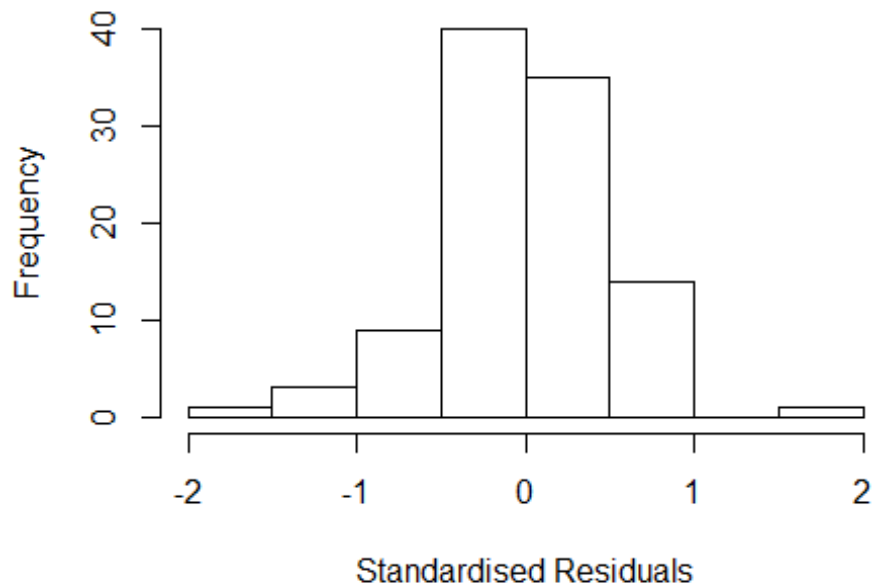


```

## Year2007          -0.9828      0.2736   -3.59  0.00055 ***
## Year2008          -0.6316      0.2492   -2.54  0.01308 *
## Year2009          -0.0341      0.2602   -0.13  0.89603
## Year2010           0.0431      0.2263    0.19  0.84954
## Year2011          -0.1735      0.2495   -0.70  0.48874
## Year2012          -0.1622      0.2050   -0.79  0.43093
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.348, Adjusted R-squared:  0.218
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.158  0.889  0.957  0.902  0.986  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      9.71e-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 2.252 1          1.501
## Year              2.252 15          1.027

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.81e+00 -2.78e-01 -3.61e-16 3.04e-01 1.74e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3389 0.2154 6.22 1.8e-08 ***
## FirstAuthorFemale1 0.0892 0.2259 0.40 0.6938
## Year1998 0.0448 0.3238 0.14 0.8904
## Year1999 0.5641 0.2154 2.62 0.0104 *
## Year2000 -0.1669 0.2154 -0.77 0.4406
## Year2001 -0.2458 0.3348 -0.73 0.4650
## Year2002 0.5635 0.2965 1.90 0.0607 .
## Year2003 0.4748 1.2783 0.37 0.7112
## Year2004 0.3086 0.2682 1.15 0.2531
## Year2005 0.0314 0.3537 0.09 0.9294
## Year2006 -0.5230 0.3421 -1.53 0.1299
## Year2007 -0.9122 0.2845 -3.21 0.0019 **
```

```

## Year2008          -0.5724      0.2454   -2.33   0.0220 *
## Year2009           0.0257      0.2808    0.09   0.9272
## Year2010           0.0739      0.2677    0.28   0.7831
## Year2011          -0.0771      0.2659   -0.29   0.7727
## Year2012          -0.1392      0.2429   -0.57   0.5681
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.332, Adjusted R-squared:  0.207
## Convergence in 33 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 94 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.130  0.898  0.961  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      9.71e-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 3.745 1          1.935
## Year              3.745 15          1.045

## [1] "List of 0 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
## -1.98101 -0.26480 -0.00799  0.30236  1.74984
##

```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3969    0.2260   6.18 2.1e-08 ***
## LastAuthorFemale1 -0.2475    0.2077  -1.19  0.2368
## Year1998          0.1096    0.2823   0.39  0.6989
## Year1999          0.5061    0.2260   2.24  0.0277 *
## Year2000         -0.2249    0.2260  -0.99  0.3226
## Year2001         -0.2712    0.3955  -0.69  0.4947
## Year2002          0.5052    0.3041   1.66  0.1002
## Year2003          0.5841    0.9298   0.63  0.5315
## Year2004          0.2782    0.2780   1.00  0.3197
## Year2005         -0.0257    0.3606  -0.07  0.9434
## Year2006         -0.5364    0.2931  -1.83  0.0707 .
## Year2007         -0.9585    0.2840  -3.38  0.0011 **
## Year2008         -0.6357    0.2534  -2.51  0.0140 *
## Year2009         -0.0259    0.2814  -0.09  0.9269
## Year2010          0.0526    0.2587   0.20  0.8394
## Year2011         -0.1061    0.2574  -0.41  0.6811
## Year2012         -0.1517    0.2379  -0.64  0.5253
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.357, Adjusted R-squared:  0.238
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 93 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0471 0.9020 0.9600 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      9.71e-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: 103"
## [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
##    2    5   10   16    4    1    7    2    7    6    1    2    6    6    9
## 2011 2012
##    3    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    9    1    1    6    2    7    5    0    2    4    5    6
## 2011 2012
##    2    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    8    1    1    6    2    7    5    0    1    4    4    6
## 2011 2012
##    2    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 3.30975091964687"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

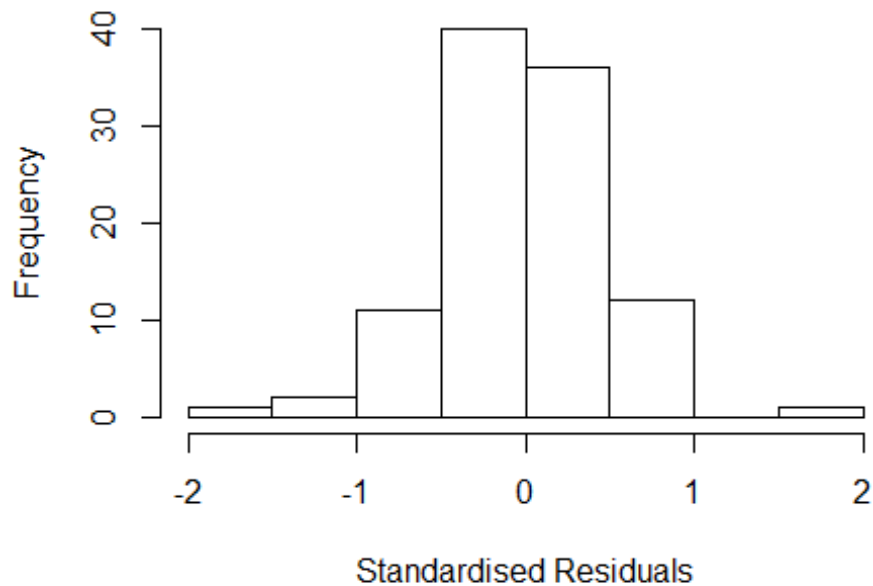
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 3.24534222319921"
## [1] "Regression 1: First author gender, last author gender, team size,
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

```

Residuals from last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN  1           NaN
## LastAuthorFemale  NaN  1           NaN
## UniqueAuthors    NaN  4           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 + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9500 -0.2785 -0.0145  0.2294  1.2342
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.6780     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1  0.6218     0.5699    1.09  0.2836
## LastAuthorFemale1 -0.3578     0.3366   -1.06  0.2961
## UniqueAuthors2    -0.2842     0.2528   -1.12  0.2696
```

```

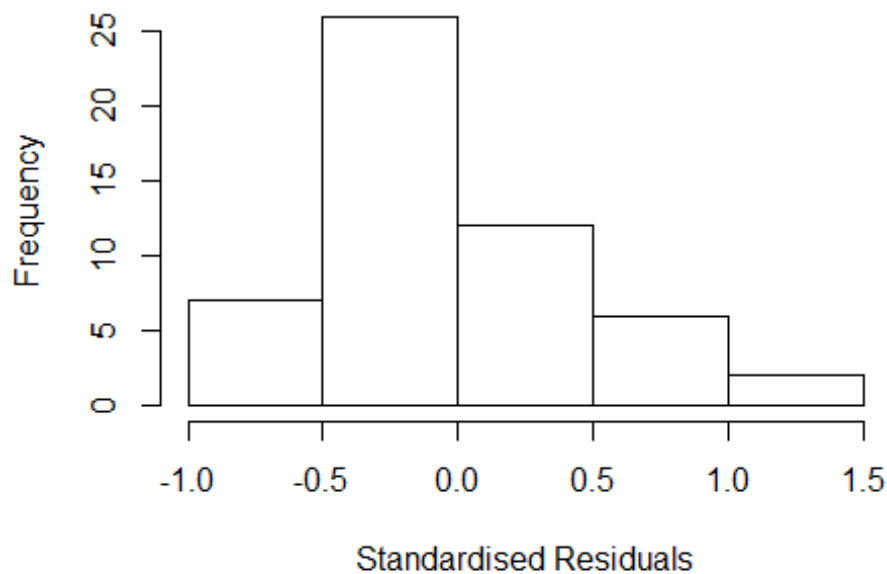
## UniqueAuthors3      -0.0612      0.2900      -0.21      0.8343
## UniqueAuthors4       0.0593      0.3395       0.17      0.8626
## UniqueAuthors5       0.4107      0.4463       0.92      0.3646
## Year1997              1.4132      0.2528       5.59      3.9e-06 ***
## Year1998             -1.3938      0.2528      -5.51      4.9e-06 ***
## Year1999             -0.7891      0.3254      -2.42      0.0213 *
## Year2000             -0.1960      0.0000      -Inf      < 2e-16 ***
## Year2001              0.7510      0.0000       Inf      < 2e-16 ***
## Year2002             -0.1244      0.2434      -0.51      0.6128
## Year2003             -0.9695      0.0714     -13.57      1.4e-14 ***
## Year2004             -0.5740      0.2419      -2.37      0.0240 *
## Year2005             -0.7118      0.2690      -2.65      0.0127 *
## Year2007             -1.6780      0.0000      -Inf      < 2e-16 ***
## Year2008             -0.5128      0.4079      -1.26      0.2181
## Year2009             -0.7838      0.3382      -2.32      0.0273 *
## Year2010             -1.3265      0.2732      -4.85      3.3e-05 ***
## Year2011             -0.0439      0.2982      -0.15      0.8838
## Year2012             -0.7020      0.2270      -3.09      0.0042 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.593
## Multiple R-squared:  0.551, Adjusted R-squared:  0.248
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.644  0.920   0.968   0.931   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      1.89e-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

```

```
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author and team size



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1          NaN
## 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 ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.64e-01 -1.93e-01 -8.88e-16  2.00e-01  1.17e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.68e+00   1.71e-08  9.83e+07 < 2e-16 ***
## FirstAuthorFemale1  7.90e-01   5.01e-01  1.58e+00  0.1237
## LastAuthorFemale1 -3.26e-01   2.83e-01 -1.15e+00  0.2575
```



```

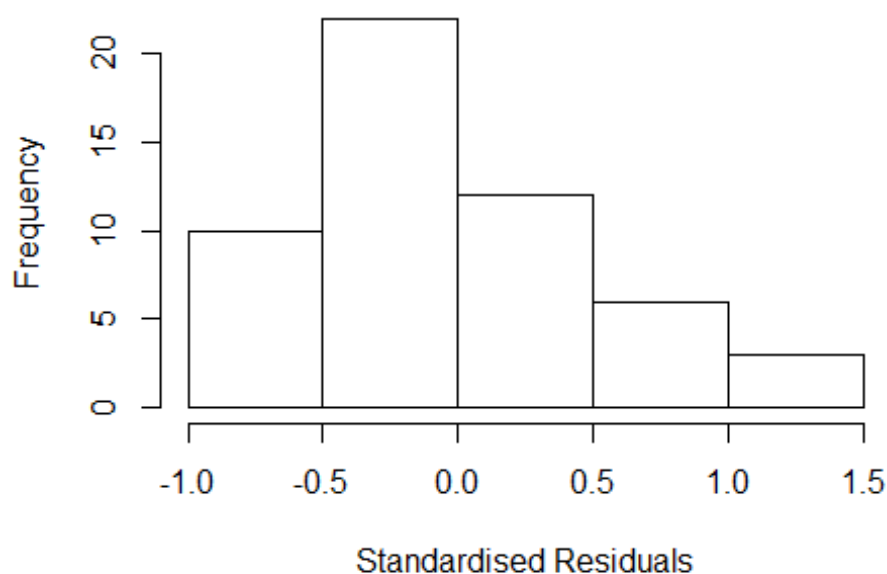
## Year1997          1.13e+00  1.76e-08  6.40e+07  < 2e-16 ***
## Year1998          -1.68e+00  2.07e-08 -8.10e+07  < 2e-16 ***
## Year1999          -7.14e-01  2.64e-01 -2.71e+00  0.0104 *
## Year2000          -1.96e-01  2.07e-08 -9.46e+06  < 2e-16 ***
## Year2001           7.51e-01  0.00e+00      Inf  < 2e-16 ***
## Year2002          -2.87e-01  1.78e-01 -1.61e+00  0.1153
## Year2003          -9.70e-01  7.15e-02 -1.36e+01  1.7e-15 ***
## Year2004          -7.35e-01  2.22e-01 -3.30e+00  0.0022 **
## Year2005          -7.77e-01  3.03e-01 -2.57e+00  0.0146 *
## Year2007          -1.68e+00  3.07e-08 -5.47e+07  < 2e-16 ***
## Year2008          -7.73e-01  3.38e-01 -2.29e+00  0.0283 *
## Year2009          -1.03e+00  4.06e-01 -2.53e+00  0.0162 *
## Year2010          -1.32e+00  1.84e-01 -7.17e+00  2.3e-08 ***
## Year2011          -2.32e-01  1.66e-01 -1.40e+00  0.1693
## Year2012          -7.85e-01  1.07e-01 -7.37e+00  1.3e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.568
## Multiple R-squared:  0.522, Adjusted R-squared:  0.289
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.653  0.888  0.962  0.920  0.994  0.999
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          1.89e-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"

## 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

```

Residuals from first and last author



```
##              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
## -0.96458 -0.28350 -0.00731  0.26735  1.19225
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.68e+00   3.23e-08  5.19e+07 < 2e-16 ***
## FirstAuthorFemale1 6.40e-01   4.37e-01  1.46e+00  0.1519
## Year1997         1.13e+00   3.42e-08  3.30e+07 < 2e-16 ***
## Year1998        -1.68e+00   0.00e+00    -Inf < 2e-16 ***
## Year1999        -7.13e-01   2.66e-01 -2.68e+00  0.0111 *
## Year2000        -1.96e-01   0.00e+00    -Inf < 2e-16 ***
## Year2001         7.51e-01   3.23e-08  2.33e+07 < 2e-16 ***
## Year2002        -3.00e-01   1.80e-01 -1.67e+00  0.1031
```

```

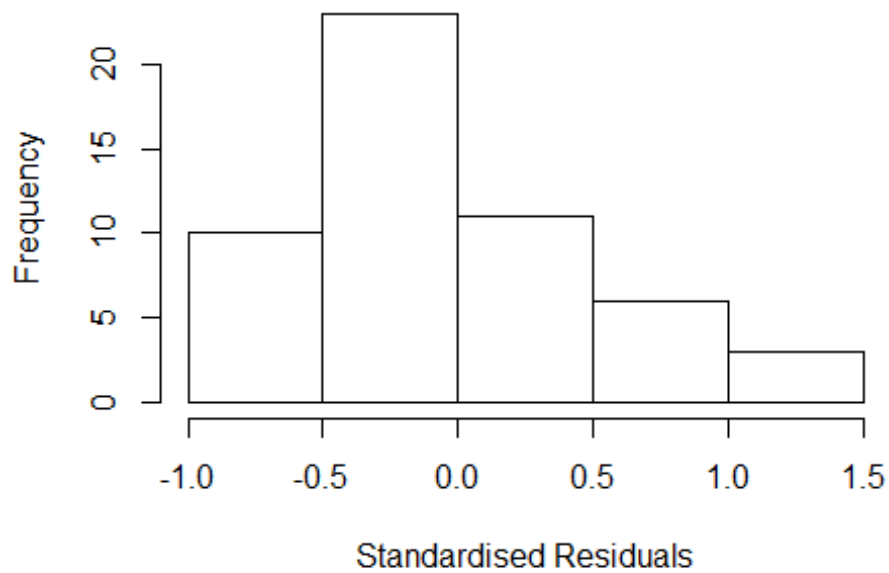
## Year2003      -9.70e-01   7.15e-02 -1.36e+01   1.0e-15 ***
## Year2004      -7.36e-01   2.24e-01 -3.29e+00   0.0022 **
## Year2005      -7.81e-01   3.04e-01 -2.57e+00   0.0144 *
## Year2007      -1.68e+00   3.04e-08 -5.53e+07   < 2e-16 ***
## Year2008      -7.77e-01   3.38e-01 -2.30e+00   0.0276 *
## Year2009      -1.05e+00   4.17e-01 -2.52e+00   0.0162 *
## Year2010      -1.30e+00   1.98e-01 -6.55e+00   1.3e-07 ***
## Year2011      -3.96e-01   2.11e-01 -1.88e+00   0.0685 .
## Year2012      -8.94e-01   5.17e-02 -1.73e+01   < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.556
## Multiple R-squared:  0.513, Adjusted R-squared:  0.297
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~ 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.625 0.890 0.958 0.913 0.991 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      1.89e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

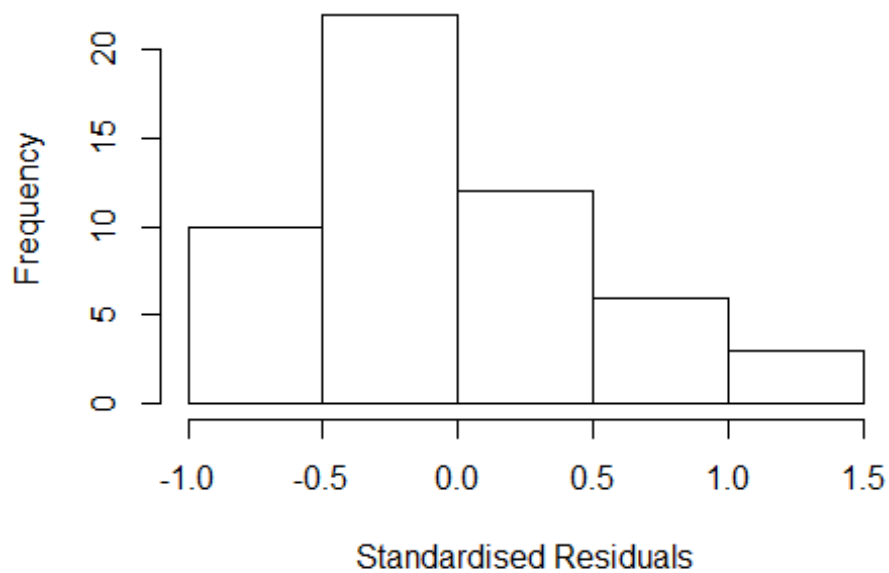
```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale NaN  1          NaN
## Year            NaN 15          NaN
```

Residuals from last author

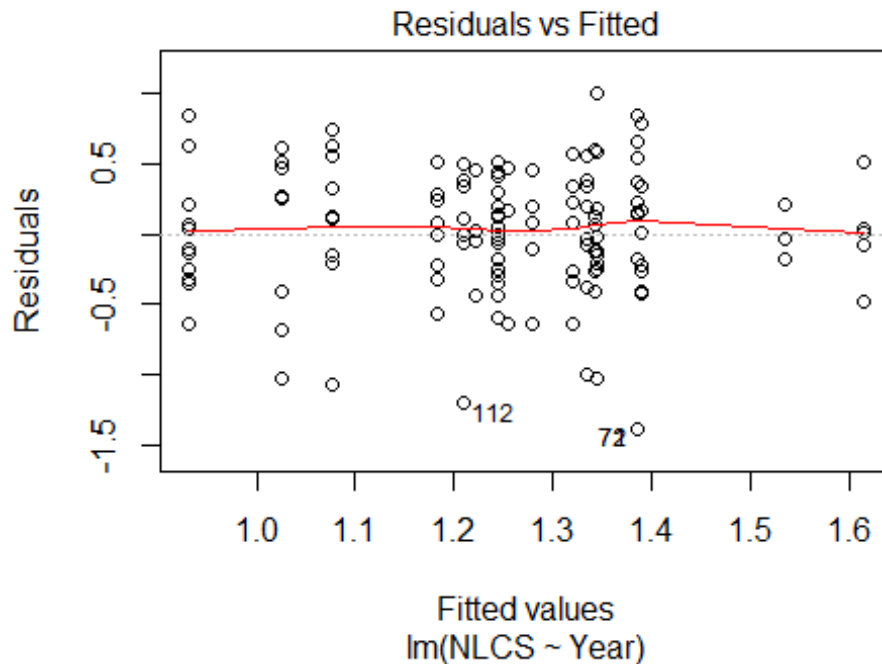


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9651 -0.2728 -0.0198 0.2709 1.7626
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.68e+00 0.00e+00 Inf < 2e-16 ***
## LastAuthorFemale1 2.17e-02 2.04e-01 1.10e-01 0.91614
## Year1997 1.13e+00 0.00e+00 Inf < 2e-16 ***
## Year1998 -1.68e+00 0.00e+00 -Inf < 2e-16 ***
## Year1999 -7.13e-01 2.68e-01 -2.66e+00 0.01161 *
## Year2000 -1.96e-01 0.00e+00 -Inf < 2e-16 ***
## Year2001 7.51e-01 1.39e-08 5.39e+07 < 2e-16 ***
## Year2002 -1.46e-01 2.88e-01 -5.10e-01 0.61451
## Year2003 -9.70e-01 7.15e-02 -1.36e+01 1.0e-15 ***
## Year2004 -7.36e-01 2.24e-01 -3.28e+00 0.00230 **
## Year2005 -7.83e-01 3.05e-01 -2.57e+00 0.01443 *
## Year2007 -1.68e+00 0.00e+00 -Inf < 2e-16 ***
## Year2008 -7.80e-01 3.39e-01 -2.30e+00 0.02733 *
## Year2009 -7.16e-01 1.94e-01 -3.70e+00 0.00072 ***
## Year2010 -1.23e+00 1.94e-01 -6.33e+00 2.5e-07 ***
## Year2011 -4.06e-01 2.42e-01 -1.68e+00 0.10228
## Year2012 -9.01e-01 8.59e-02 -1.05e+01 1.7e-12 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared: 0.489, Adjusted R-squared: 0.262
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~ 1. The remaining 41 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.281 0.887 0.958 0.901 0.992 0.999
## Algorithmic parameters:
## tuning.chi bb tuning.psi refine.tol
## 1.55e+00 5.00e-01 4.69e+00 1.00e-07
## rel.tol solve.tol eps.outlier eps.x
## 1.00e-07 1.00e-07 1.89e-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: 53"
## [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
##    4    9   12    6   10   14   15   13    8    8   11   14   12   13   28
## 2011 2012
##   12   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    6    8    3    3   10    5   10    8    5    8   11    8    9   18
## 2011 2012
##    8    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    5    6    3    2    9    4    7    8    4    8   11    8    8   18
## 2011 2012
##    8    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##      Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 19, df = 16, p-value = 0.3

```



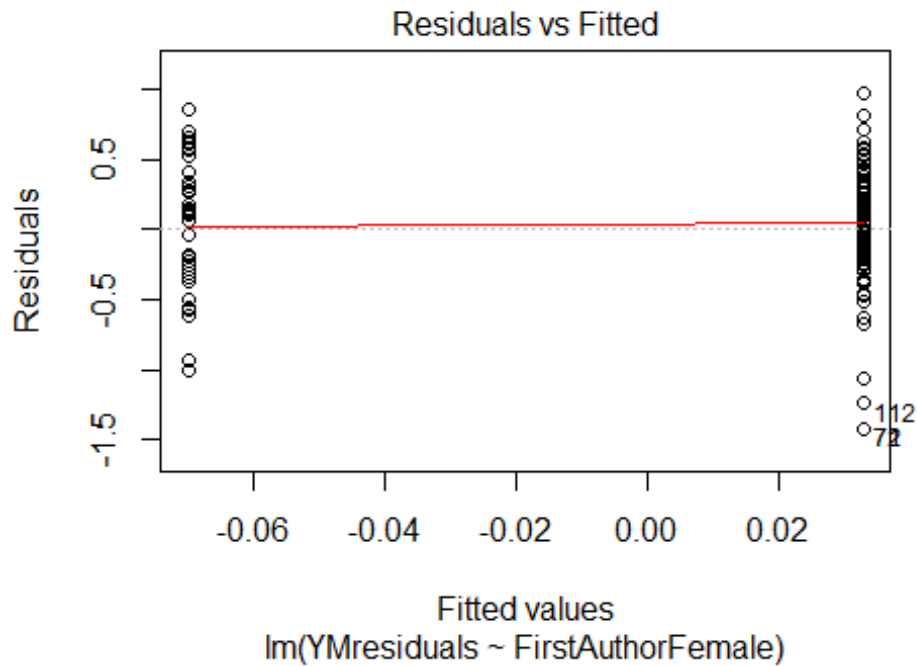
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.39, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 4.227335859913"
## [1] "Male first author team size 2018 geometric mean: 3.41647290008184"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

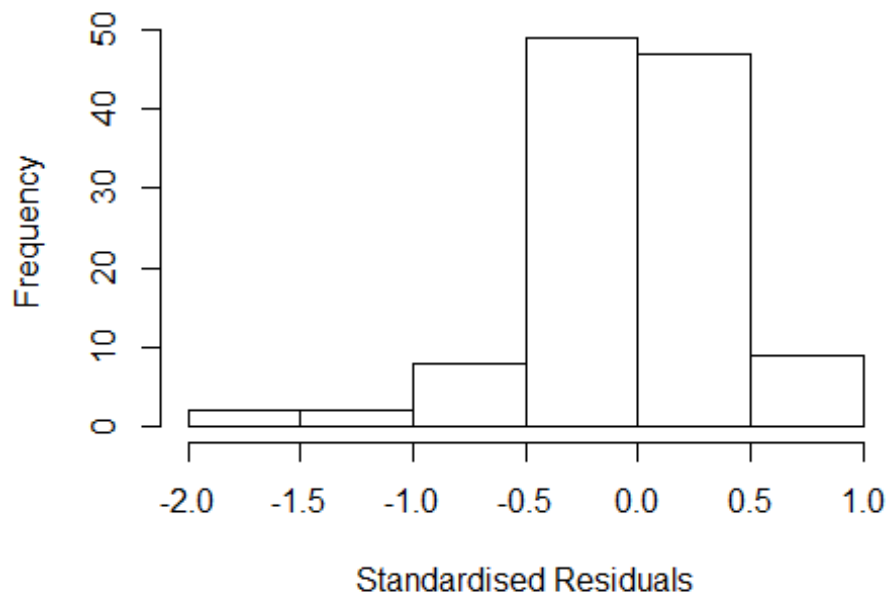
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 42, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.66284150148471"
## [1] "Male last author team size 2018 geometric mean: 3.61384601356524"

## 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 = 26, 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  2.613 1          1.617
## UniqueAuthors    13.598 4          1.386
## Year              24.044 16         1.104
```


Residuals from first and last author and team size



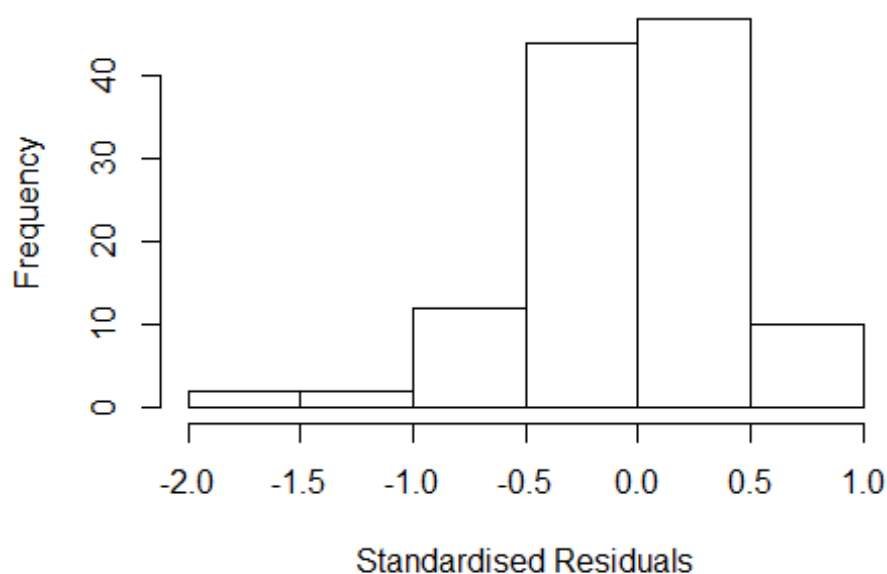
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.7600 -0.2684 -0.0415 0.2980 0.8986
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.24835 0.21808 5.72 1.2e-07 ***
## FirstAuthorFemale1 -0.15530 0.12032 -1.29 0.20
## LastAuthorFemale1 -0.10659 0.12711 -0.84 0.40
## UniqueAuthors2 0.01929 0.10799 0.18 0.86
## UniqueAuthors3 0.16353 0.12548 1.30 0.20
## UniqueAuthors4 0.17249 0.24814 0.70 0.49
## UniqueAuthors5 0.13378 0.15444 0.87 0.39
## Year1997 0.04076 0.24641 0.17 0.87
## Year1998 0.25664 0.27635 0.93 0.36
## Year1999 0.02983 0.38755 0.08 0.94
```

```

## Year2000      0.34058    0.27040    1.26    0.21
## Year2001      0.33919    0.28918    1.17    0.24
## Year2002      0.31007    0.29650    1.05    0.30
## Year2003     -0.17975    0.37331   -0.48    0.63
## Year2004      0.08870    0.26331    0.34    0.74
## Year2005      0.06742    0.31162    0.22    0.83
## Year2006     -0.01099    0.25644   -0.04    0.97
## Year2007     -0.36590    0.24242   -1.51    0.13
## Year2008     -0.16057    0.31081   -0.52    0.61
## Year2009      0.20458    0.26564    0.77    0.44
## Year2010      0.00213    0.21924    0.01    0.99
## Year2011      0.12613    0.27385    0.46    0.65
## Year2012     -0.02060    0.30231   -0.07    0.95
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.435
## Multiple R-squared:  0.22,   Adjusted R-squared:  0.037
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0639 0.9060 0.9560 0.9030 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      8.55e-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 1.732 1      1.316
## LastAuthorFemale 2.170 1      1.473
## Year      2.879 16      1.034

```

Residuals from first and last author



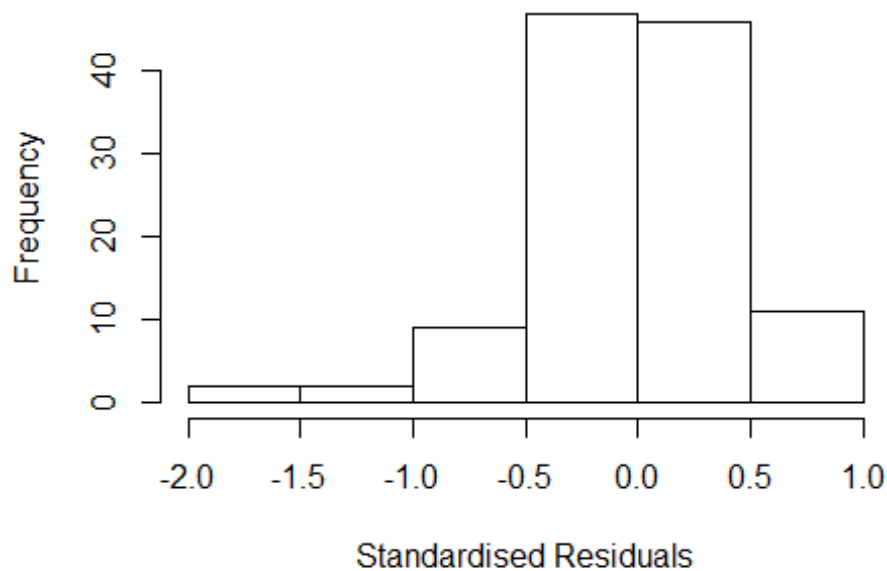
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5997 -0.2534 -0.0218 0.2568 0.8516
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2914 0.1960 6.59 2.2e-09 ***
## FirstAuthorFemale1 -0.1156 0.1147 -1.01 0.316
## LastAuthorFemale1 -0.1293 0.1248 -1.04 0.302
## Year1997 0.0430 0.2487 0.17 0.863
## Year1998 0.2535 0.2933 0.86 0.389
## Year1999 0.0200 0.3415 0.06 0.954
## Year2000 0.4571 0.2276 2.01 0.047 *
## Year2001 0.3083 0.2777 1.11 0.270
## Year2002 0.3449 0.3176 1.09 0.280
## Year2003 -0.1442 0.3905 -0.37 0.713
## Year2004 0.0975 0.2474 0.39 0.694
## Year2005 0.0656 0.2822 0.23 0.817
```

```

## Year2006          -0.0313      0.2366   -0.13    0.895
## Year2007          -0.3620      0.2334   -1.55    0.124
## Year2008          -0.0692      0.3252   -0.21    0.832
## Year2009           0.2064      0.2559    0.81    0.422
## Year2010           0.0189      0.2130    0.09    0.930
## Year2011           0.1830      0.2584    0.71    0.480
## Year2012           0.0493      0.2979    0.17    0.869
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.192, Adjusted R-squared:  0.0432
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.150 0.898 0.960 0.902 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      8.55e-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 1.687 1      1.299
## Year              1.687 16      1.016

```

Residuals from first author



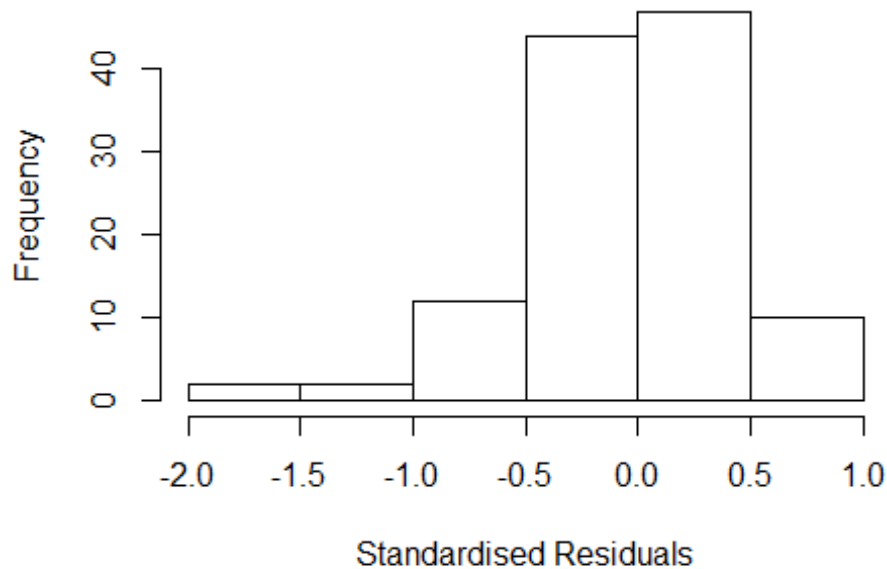
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6058 -0.2624 -0.0369 0.2633 0.8800
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.28787 0.17200 7.49 3e-11 ***
## FirstAuthorFemale1 -0.15824 0.11822 -1.34 0.184
## Year1997 0.05550 0.22654 0.25 0.807
## Year1998 0.24401 0.27006 0.90 0.368
## Year1999 0.03394 0.31419 0.11 0.914
## Year2000 0.41725 0.22734 1.84 0.069 .
## Year2001 0.31792 0.26148 1.22 0.227
## Year2002 0.33776 0.28598 1.18 0.240
## Year2003 -0.13327 0.41937 -0.32 0.751
## Year2004 0.06959 0.22319 0.31 0.756
## Year2005 0.02749 0.26154 0.11 0.917
## Year2006 -0.02237 0.20819 -0.11 0.915
```

```

## Year2007          -0.34897    0.21313   -1.64    0.105
## Year2008          -0.11283    0.28651   -0.39    0.695
## Year2009           0.18015    0.23436    0.77    0.444
## Year2010          -0.00417    0.18501   -0.02    0.982
## Year2011           0.14333    0.23252    0.62    0.539
## Year2012          -0.03982    0.28222   -0.14    0.888
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.437
## Multiple R-squared:  0.181, Adjusted R-squared:  0.0402
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 110 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.148  0.892  0.960  0.903  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      8.55e-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 2.781 1          1.668
## Year              2.781 16          1.032

```

Residuals from last author



```
## [1] "List of 0 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
## -1.5860 -0.2299 -0.0289 0.2604 0.8777
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2544 0.2052 6.11 2e-08 ***
## LastAuthorFemale1 -0.1781 0.1267 -1.41 0.163
## Year1997 0.0565 0.2647 0.21 0.831
## Year1998 0.2696 0.2816 0.96 0.341
## Year1999 0.0325 0.3852 0.08 0.933
## Year2000 0.4607 0.2204 2.09 0.039 *
## Year2001 0.3316 0.2753 1.20 0.231
## Year2002 0.3735 0.3095 1.21 0.230
## Year2003 -0.1193 0.3968 -0.30 0.764
## Year2004 0.1180 0.2463 0.48 0.633
## Year2005 0.0707 0.2906 0.24 0.808
## Year2006 -0.0435 0.2475 -0.18 0.861
```

```

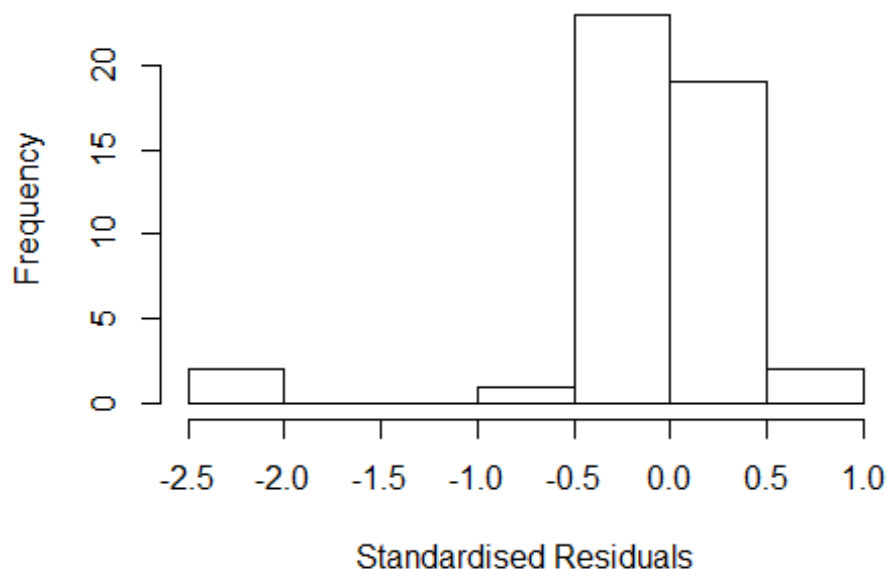
## Year2007          -0.3510      0.2414    -1.45      0.149
## Year2008          -0.0506      0.3476    -0.15      0.885
## Year2009           0.2275      0.2557      0.89      0.376
## Year2010           0.0415      0.2221      0.19      0.852
## Year2011           0.2082      0.2677      0.78      0.439
## Year2012           0.1210      0.3007      0.40      0.688
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.432
## Multiple R-squared:  0.181, Adjusted R-squared:  0.04
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 107 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.150  0.883   0.956   0.899   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      8.55e-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: 117"
## [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
##    7    9   14    8    7   12    5    3    1    5    1    6    1    4    1
## 2012
##    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6   10    5    4    4    3    1    1    4    0    6    0    4    0
## 2012

```



```
##      1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      4      6      9      4      4      3      3      0      1      4      0      4      0      4      0
## 2012
##      1
## [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: 1.44224957030741"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -7.053e+14  1      NaN
## LastAuthorFemale  -6.270e+00  1      NaN
## UniqueAuthors    -8.736e+30  4      NaN
## Year              -1.769e+32 11      NaN
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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,
```

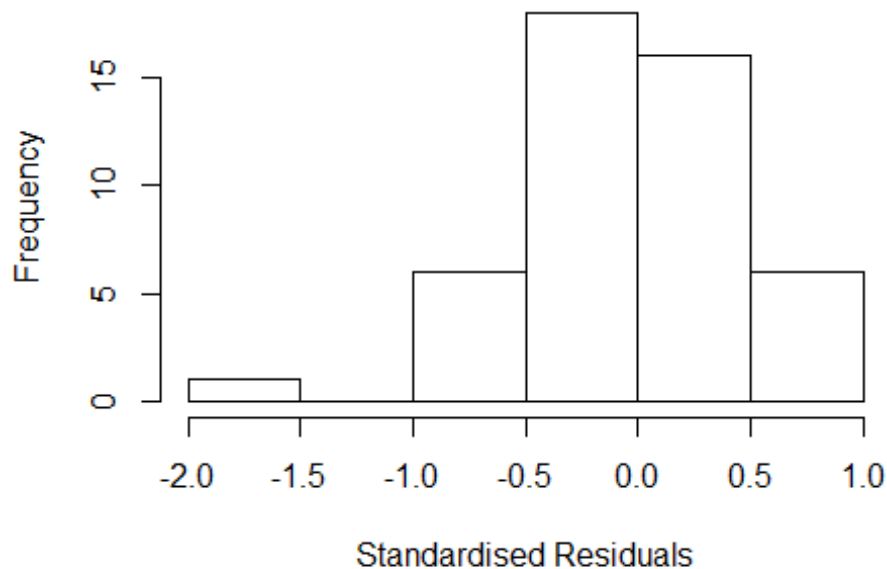
```

##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q        Max
## -2.33e+00 -2.57e-01 -4.86e-16  1.53e-01  6.62e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1755     0.2041    5.76 3.1e-06 ***
## FirstAuthorFemale1  0.9124     0.1722    5.30 1.1e-05 ***
## LastAuthorFemale1 -0.4249     0.1431   -2.97 0.00592 **
## UniqueAuthors2    -0.4290     0.1412   -3.04 0.00500 **
## UniqueAuthors3    -0.7233     0.1236   -5.85 2.4e-06 ***
## UniqueAuthors4     0.3110     0.1683    1.85 0.07482 .
## UniqueAuthors5    -0.3375     0.1041   -3.24 0.00297 **
## Year1997           0.4621     0.2769    1.67 0.10595
## Year1998           0.0619     0.2535    0.24 0.80886
## Year1999           1.0867     0.3546    3.07 0.00467 **
## Year2000           0.6969     0.2507    2.78 0.00944 **
## Year2001           0.2649     0.2353    1.13 0.26942
## Year2002          -0.7253     0.3415   -2.12 0.04232 *
## Year2004           0.2295     0.2041    1.12 0.27012
## Year2005           0.3520     0.2279    1.54 0.13327
## Year2007          -0.3742     0.2856   -1.31 0.20040
## Year2009          -1.1774     0.2717   -4.33 0.00016 ***
## Year2012           0.1034     0.3049    0.34 0.73705
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.757, Adjusted R-squared:  0.614
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 2 observations c(10,20) are outliers with |weight| = 0 ( < 0.0021);
## 8 weights are ~ = 1. The remaining 37 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.787  0.944  0.983  0.958  0.993  0.999
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          2.13e-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 1.104e+15 1 3.322e+07
## LastAuthorFemale 4.722e+00 1 2.173e+00
## Year 5.628e+15 11 5.199e+00
```

Residuals from first and last author



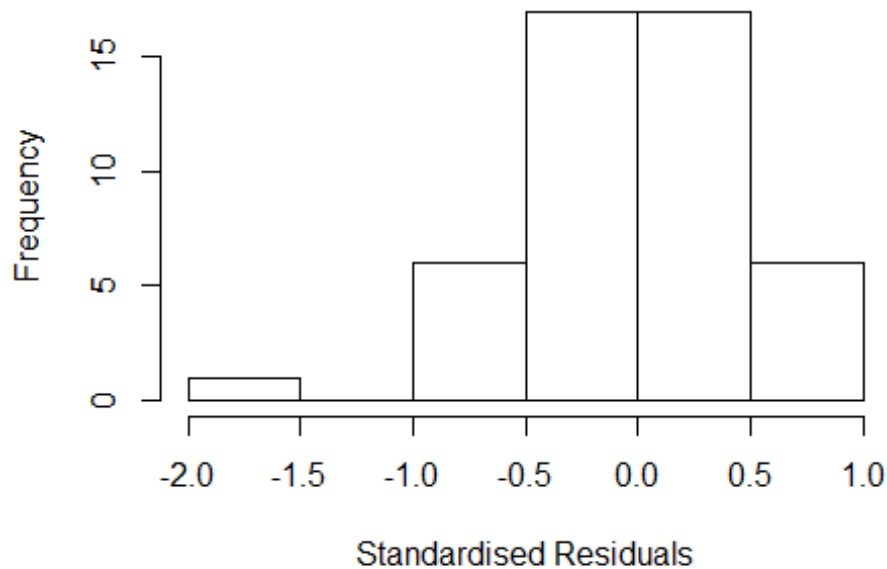
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.65e+00 -3.23e-01 -1.39e-16 3.57e-01 8.57e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.175500 0.196546 5.98 1e-06 ***
```

```

## FirstAuthorFemale1  0.293289    0.297961    0.98  0.33212
## LastAuthorFemale1   0.142100    0.243235    0.58  0.56305
## Year1997            -0.017751    0.322267   -0.06  0.95640
## Year1998            -0.044913    0.276442   -0.16  0.87193
## Year1999             0.474428    0.500340    0.95  0.34991
## Year2000             0.445972    0.251556    1.77  0.08548 .
## Year2001            -0.022008    0.260299   -0.08  0.93313
## Year2002            -1.011630    0.241072   -4.20  0.00019 ***
## Year2004             0.229500    0.196546    1.17  0.25131
## Year2005            -0.015230    0.268923   -0.06  0.95518
## Year2007            -0.489160    0.391605   -1.25  0.22041
## Year2009            -1.284339    0.467867   -2.75  0.00971 **
## Year2012            -0.000789    0.356947    0.00  0.99825
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.56
## Multiple R-squared:  0.478, Adjusted R-squared:  0.272
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.366  0.914  0.962  0.930  0.992  0.999
## Algorithmic parameters:
##           tuning.chi                bb           tuning.psi           refine.tol
##           1.55e+00                5.00e-01           4.69e+00           1.00e-07
##           rel.tol                solve.tol           eps.outlier           eps.x
##           1.00e-07                1.00e-07           2.13e-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 6.469e+14 1         2.543e+07
## Year              6.469e+14 11         4.712e+00

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.66e+00 -3.29e-01 1.09e-15 3.44e-01 8.34e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.17550 0.19684 5.97 9.4e-07 ***
## FirstAuthorFemale1 0.32529 0.28256 1.15 0.25766
## Year1997 0.00515 0.30769 0.02 0.98675
## Year1998 -0.04929 0.27657 -0.18 0.85962
## Year1999 0.48049 0.49576 0.97 0.33929
## Year2000 0.48354 0.23762 2.03 0.04972 *
## Year2001 -0.02201 0.26059 -0.08 0.93318
## Year2002 -1.01168 0.24134 -4.19 0.00019 ***
## Year2004 0.22950 0.19684 1.17 0.25175
## Year2005 -0.01512 0.26919 -0.06 0.95553
## Year2007 -0.47226 0.39964 -1.18 0.24552
## Year2009 -1.19739 0.37059 -3.23 0.00274 **
```

```

## Year2012          -0.03279    0.34436   -0.10  0.92469
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.555
## Multiple R-squared:  0.478, Adjusted R-squared:  0.294
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.354  0.914  0.964  0.930  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      2.13e-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 7.33  1         2.707
## Year              7.33 11         1.095
##
## [1] "List of 0 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
## -1.63e+00 -3.13e-01 -2.22e-16  3.57e-01  7.92e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.17550    0.19552     6.01  8.3e-07 ***
## LastAuthorFemale1 0.21848    0.23865     0.92  0.36639

```

```

## Year1997      0.04746      0.32809      0.14  0.88583
## Year1998     -0.00547      0.26993     -0.02  0.98397
## Year1999      0.45320      0.51023      0.89  0.38066
## Year2000      0.42606      0.25028      1.70  0.09782 .
## Year2001     -0.02200      0.25927     -0.08  0.93288
## Year2002     -1.01145      0.24010     -4.21  0.00018 ***
## Year2004      0.22950      0.19552      1.17  0.24864
## Year2005     -0.01561      0.26783     -0.06  0.95387
## Year2007     -0.35838      0.45521     -0.79  0.43656
## Year2009     -1.11419      0.33033     -3.37  0.00187 **
## Year2012      0.29250      0.19552      1.50  0.14388
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.578
## Multiple R-squared:  0.455, Adjusted R-squared:  0.262
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.407  0.902  0.964  0.934  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      2.13e-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: 47"
## [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
##    4    4    9   12    2    9    8    4    4    6    9    9   21   16   20
## 2011 2012

```

```

## 29 25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 2 2 7 0 3 2 1 1 4 5 4 9 8 14
## 2011 2012
## 17 15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 0 2 2 6 0 3 2 1 1 3 5 2 8 7 14
## 2011 2012
## 13 11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.7062880231628"
## [1] "Male first author team size 2018 geometric mean: 3.01774902754934"

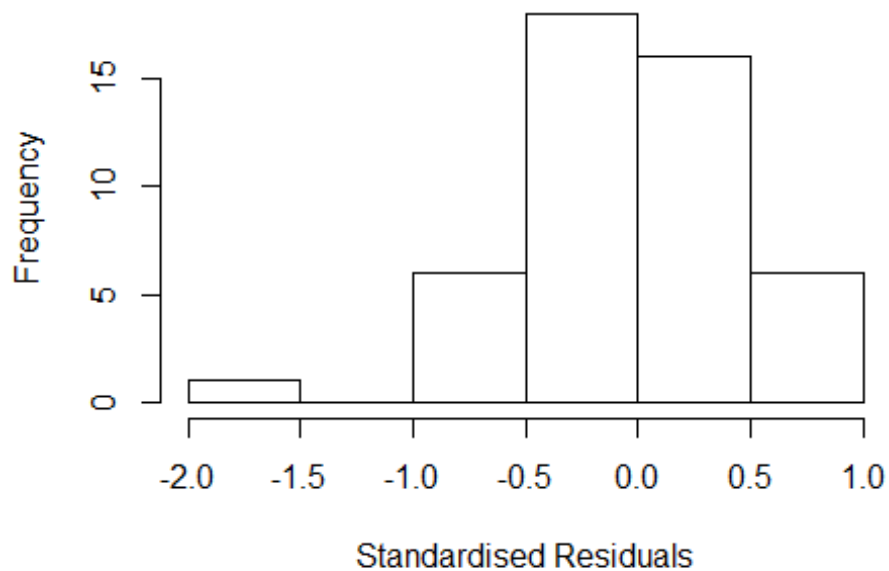
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 58, p-value = 0.3
## 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: 3.27636734740592"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

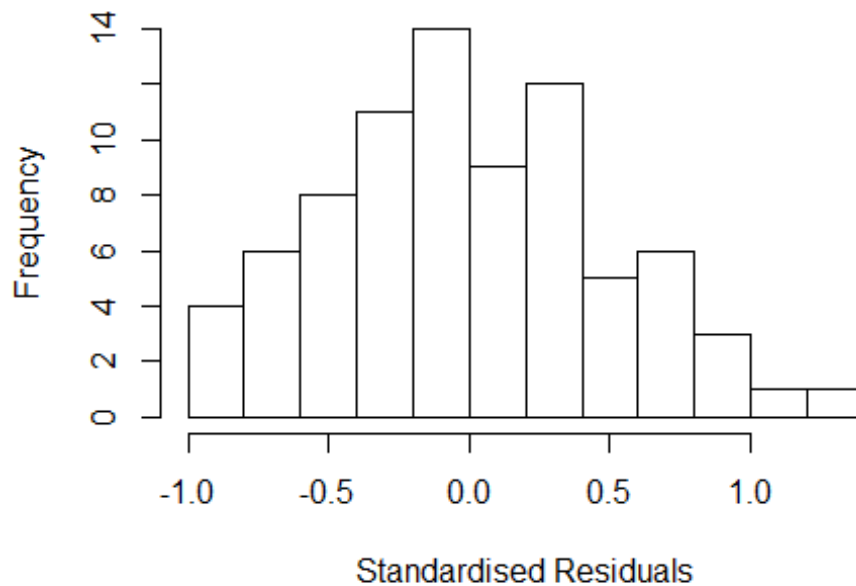
```


Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2.5, 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 3.532e+00 1          1.879
## LastAuthorFemale  3.530e+00 1          1.879
## UniqueAuthors    1.908e+16 4         108.412
## Year              3.541e+16 14          3.900
```

Residuals from first and last author and team size



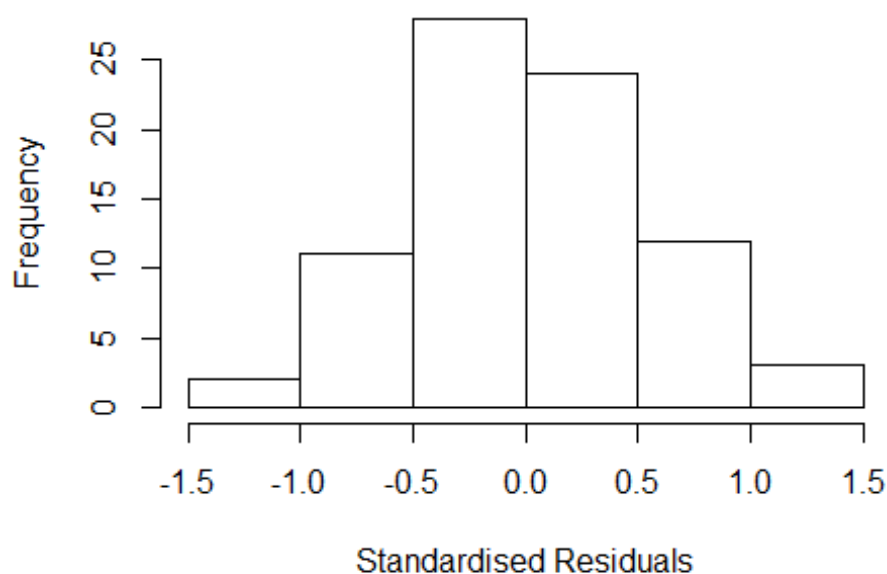
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9169 -0.3493 -0.0212 0.3500 1.2406
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.83150 0.18203 4.57 2.6e-05 ***
## FirstAuthorFemale1 0.02353 0.48718 0.05 0.9616
## LastAuthorFemale1 -0.08641 0.36738 -0.24 0.8149
## UniqueAuthors2 -0.02726 0.24694 -0.11 0.9125
## UniqueAuthors3 -0.32939 0.25787 -1.28 0.2065
## UniqueAuthors4 -0.40060 0.35229 -1.14 0.2601
## UniqueAuthors5 -0.05544 0.29854 -0.19 0.8533
## Year1998 0.26050 0.26104 1.00 0.3224
## Year1999 0.19396 0.27896 0.70 0.4896
## Year2001 0.00633 0.45485 0.01 0.9889
```

```

## Year2002          0.47820      0.38931      1.23      0.2242
## Year2003          0.05389      0.31565      0.17      0.8650
## Year2004         -0.39924      0.30678     -1.30      0.1982
## Year2005         -0.21406      0.68920     -0.31      0.7572
## Year2006          0.21382      0.36400      0.59      0.5592
## Year2007          0.52950      0.33197      1.60      0.1161
## Year2008          0.27678      0.32531      0.85      0.3983
## Year2009         -0.04190      0.29985     -0.14      0.8894
## Year2010          0.90471      0.33222      2.72      0.0085 **
## Year2011          0.21911      0.29428      0.74      0.4595
## Year2012          0.36166      0.35314      1.02      0.3100
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.609
## Multiple R-squared:  0.252, Adjusted R-squared:  -0.00151
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.658  0.895   0.961   0.933   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      1.25e-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 -2.158e+10  1      NaN
## LastAuthorFemale  -1.918e+10  1      NaN
## Year              -1.558e+10 14      NaN

```

Residuals from first and last author



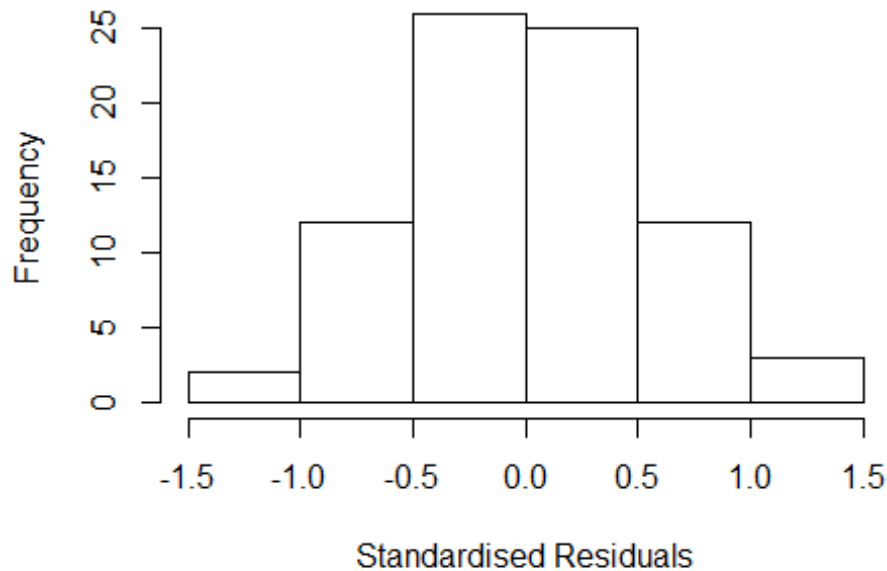
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.01e+00 -3.94e-01 3.05e-16 3.00e-01 1.27e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.83150 0.18247 4.56 2.4e-05 ***
## FirstAuthorFemale1 0.27204 0.46509 0.58 0.5607
## LastAuthorFemale1 -0.33817 0.31719 -1.07 0.2904
## Year1998 0.26050 0.26168 1.00 0.3233
## Year1999 0.19035 0.28119 0.68 0.5009
## Year2001 0.00117 0.47815 0.00 0.9980
## Year2002 0.31350 0.53343 0.59 0.5588
## Year2003 -0.27550 0.18247 -1.51 0.1361
## Year2004 -0.42650 0.18247 -2.34 0.0226 *
## Year2005 -0.24440 0.64262 -0.38 0.7050
## Year2006 0.12514 0.32929 0.38 0.7052
## Year2007 0.16450 0.18252 0.90 0.3709
```

```

## Year2008          0.18038    0.25761    0.70    0.4864
## Year2009         -0.10352    0.22426   -0.46    0.6459
## Year2010          0.73987    0.23667    3.13    0.0027 **
## Year2011          0.15965    0.24857    0.64    0.5230
## Year2012          0.18260    0.27381    0.67    0.5073
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.217, Adjusted R-squared:  0.0185
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.619  0.885   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      1.25e-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 1.652 1          1.285
## Year              1.652 14          1.018

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.01431 -0.38211 0.00279 0.32784 1.27359
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.83150 0.18255 4.55 2.4e-05 ***
## FirstAuthorFemale1 -0.04813 0.36633 -0.13 0.8959
## Year1998 0.26050 0.26179 1.00 0.3235
## Year1999 0.18664 0.28083 0.66 0.5087
## Year2001 0.00145 0.47925 0.00 0.9976
## Year2002 0.31350 0.53478 0.59 0.5598
## Year2003 -0.27550 0.18255 -1.51 0.1362
## Year2004 -0.42650 0.18255 -2.34 0.0226 *
## Year2005 -0.24709 0.64449 -0.38 0.7027
## Year2006 0.12475 0.32958 0.38 0.7063
## Year2007 0.16450 0.18260 0.90 0.3710
## Year2008 0.18066 0.25771 0.70 0.4858
```

```

## Year2009          -0.14509    0.24065   -0.60    0.5487
## Year2010          0.70991    0.23105    3.07    0.0031 **
## Year2011          0.15870    0.24828    0.64    0.5250
## Year2012          0.18281    0.27418    0.67    0.5073
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.584
## Multiple R-squared:  0.208, Adjusted R-squared:  0.0228
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.614  0.881  0.945  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      1.25e-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 2.013  1          1.419
## Year            2.013 14          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
## -1.015 -0.397  0.003  0.296  1.280
##
## Coefficients:

```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.83150    0.18273   4.55 2.4e-05 ***
## LastAuthorFemale1 -0.17272    0.28466  -0.61  0.5461
## Year1998          0.26050    0.26206   0.99  0.3239
## Year1999          0.21050    0.27503   0.77  0.4469
## Year2001          0.00211    0.48185   0.00  0.9965
## Year2002          0.31350    0.53805   0.58  0.5622
## Year2003         -0.27550    0.18273  -1.51  0.1365
## Year2004         -0.42650    0.18273  -2.33  0.0227 *
## Year2005         -0.25373    0.64725  -0.39  0.6963
## Year2006          0.12380    0.33019   0.37  0.7089
## Year2007          0.16450    0.18278   0.90  0.3715
## Year2008          0.18135    0.25790   0.70  0.4845
## Year2009         -0.12533    0.23017  -0.54  0.5880
## Year2010          0.74356    0.23836   3.12  0.0027 **
## Year2011          0.15924    0.25326   0.63  0.5317
## Year2012          0.18332    0.27500   0.67  0.5074
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.576
## Multiple R-squared:  0.217, Adjusted R-squared:  0.0332
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 66 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.600  0.880  0.944  0.917  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           1.25e-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: 80"
## [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]"
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    1    2    4    2    1    1    3    4    1    4    3    3    5    2
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    2    1    1    1    1    3    3    1    2    0    1    5    1
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    2    1    1    1    1    3    3    1    2    0    1    5    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.10723250595386"
## [1] "Male first author team size 2018 geometric mean: 1.41421356237309"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

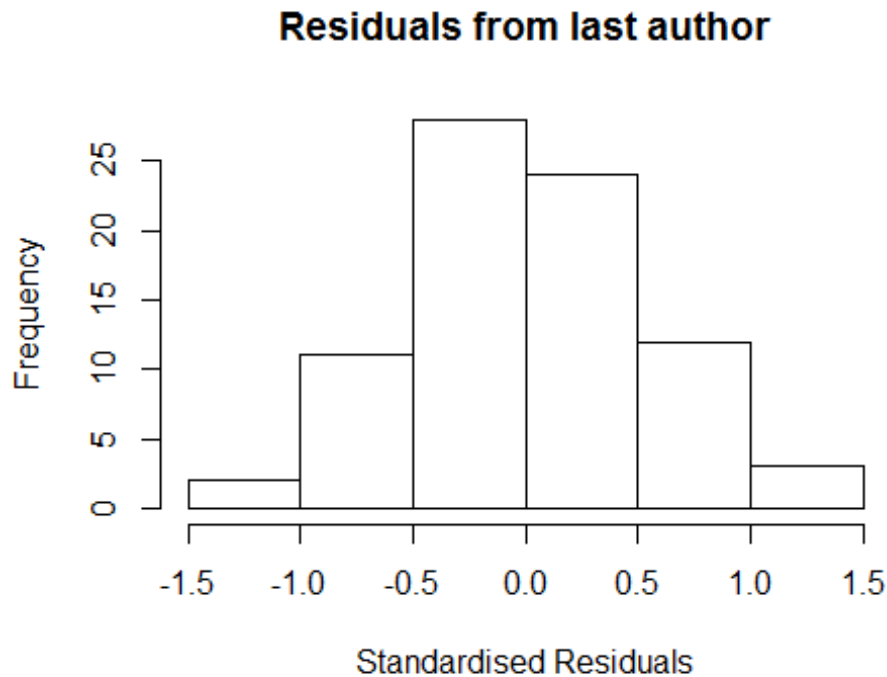
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5.5, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 2.34034731932072"

## 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 = 1.5, 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"

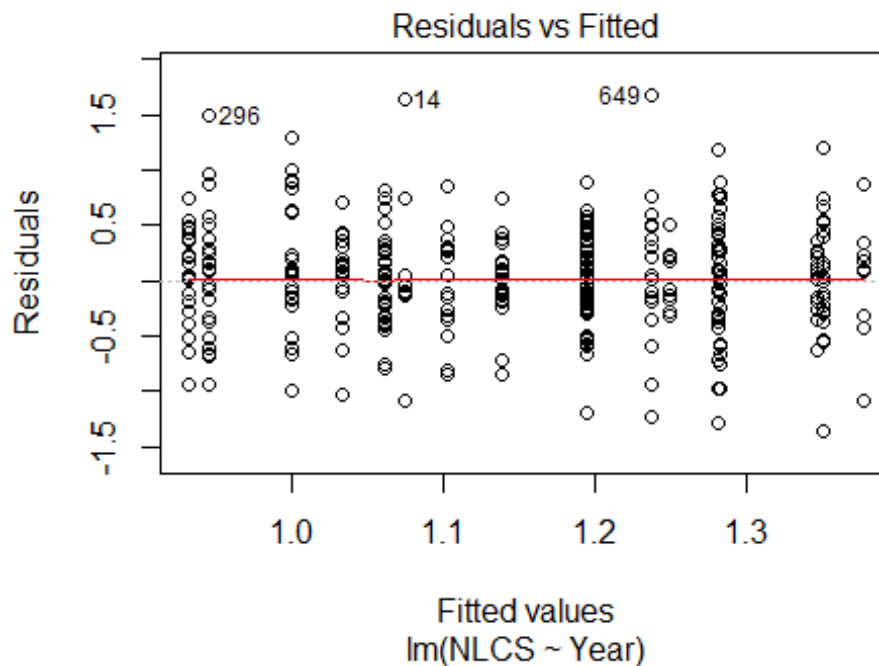
## Warning in lmrob.S(x, y, control = control, mf = mf): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

```



```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
## [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: 22"
## [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
##    26    30    21    16    37    41    33    34    38    45    40    49    50    42    28
## 2011 2012
##    37    39
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    11    19     9    10    16    17    22    25    26    32    23    34    27    29    19
## 2011 2012
##    27    24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    11    17     8    10    16    16    21    24    23    29    21    32    23    26    19
## 2011 2012
##    25    23
```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 44, df = 16, p-value = 2e-04
```



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

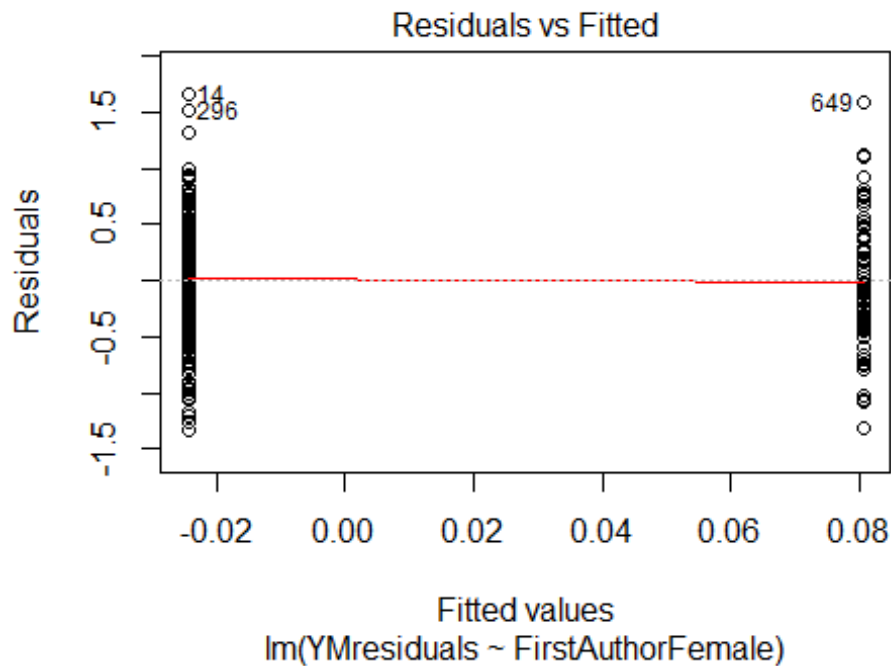
## [1] "Female first author team size 2018 geometric mean: 2.24531414385029"
## [1] "Male first author team size 2018 geometric mean: 2.34260241422694"

## 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 = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
```

```
## [1] "Female last author team size 2018 geometric mean: 2.42556498714542"
## [1] "Male last author team size 2018 geometric mean: 2.23695917152789"

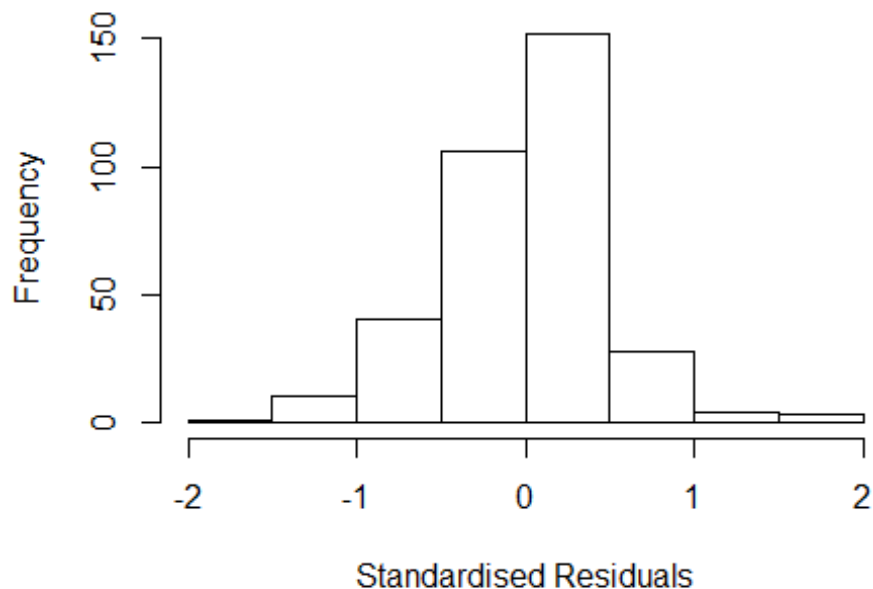
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 140, p-value = 0.5
## 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	2.074	1	1.440
LastAuthorFemale	2.113	1	1.454
UniqueAuthors	2.837	4	1.139
Year	4.482	16	1.048

Residuals from first and last author and team size



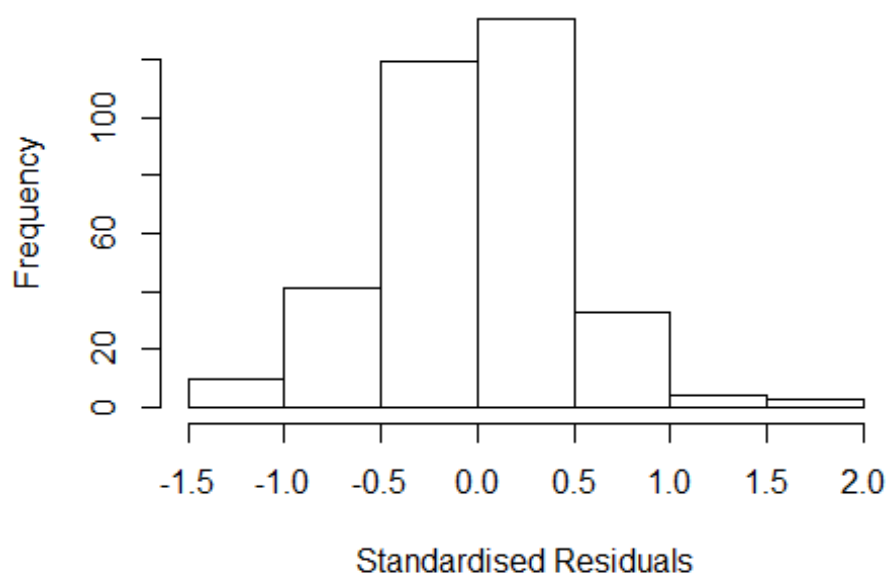
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6062 -0.2753 0.0338 0.2774 1.8405
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.87048 0.17322 5.03 8.4e-07 ***
## FirstAuthorFemale1 0.13983 0.08159 1.71 0.0875 .
## LastAuthorFemale1 -0.13226 0.07813 -1.69 0.0915 .
## UniqueAuthors2 0.16359 0.07673 2.13 0.0338 *
## UniqueAuthors3 0.33570 0.07118 4.72 3.6e-06 ***
## UniqueAuthors4 0.38516 0.08186 4.71 3.8e-06 ***
## UniqueAuthors5 0.28502 0.08719 3.27 0.0012 **
## Year1997 0.13909 0.20329 0.68 0.4943
## Year1998 0.41274 0.22397 1.84 0.0663 .
## Year1999 0.24363 0.18489 1.32 0.1886
```

```

## Year2000          0.28725    0.18716    1.53    0.1258
## Year2001          0.06233    0.21075    0.30    0.7676
## Year2002          0.03560    0.19128    0.19    0.8525
## Year2003         -0.11399    0.20954   -0.54    0.5868
## Year2004         -0.09159    0.19280   -0.48    0.6351
## Year2005         -0.00709    0.18011   -0.04    0.9686
## Year2006          0.11075    0.18027    0.61    0.5394
## Year2007          0.08020    0.18315    0.44    0.6618
## Year2008          0.11180    0.19048    0.59    0.5576
## Year2009          0.24782    0.19651    1.26    0.2082
## Year2010          0.26023    0.22922    1.14    0.2571
## Year2011         -0.12228    0.25840   -0.47    0.6364
## Year2012          0.26012    0.18745    1.39    0.1662
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.428
## Multiple R-squared:  0.203, Adjusted R-squared:  0.149
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 306 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0246 0.8620 0.9540 0.8870 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          2.91e-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 1.863 1          1.365
## LastAuthorFemale  1.842 1          1.357
## Year              1.519 16          1.013

```

Residuals from first and last author



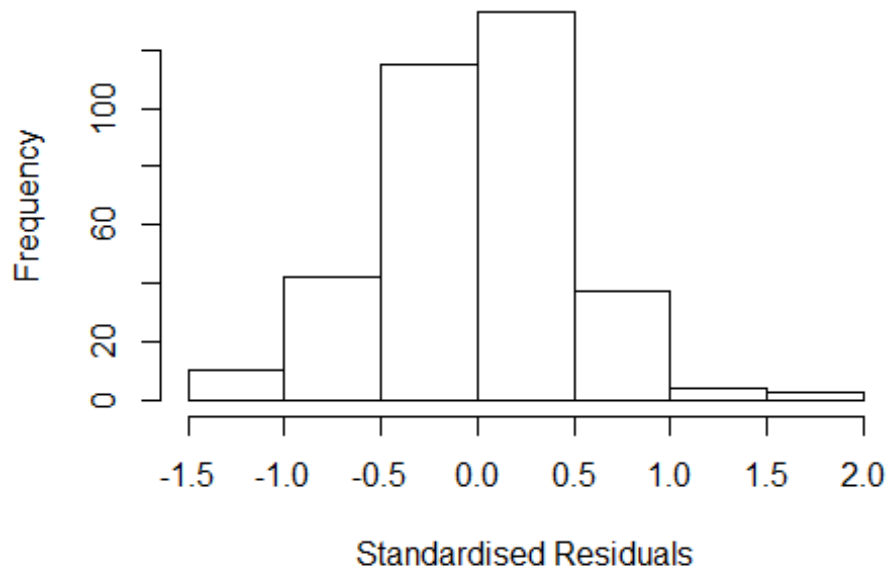
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.38755 -0.31176 0.00999 0.30229 1.73753
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.97347 0.17681 5.51 7.5e-08 ***
## FirstAuthorFemale1 0.16120 0.08620 1.87 0.062 .
## LastAuthorFemale1 -0.12244 0.08251 -1.48 0.139
## Year1997 0.20782 0.21211 0.98 0.328
## Year1998 0.42093 0.25975 1.62 0.106
## Year1999 0.27577 0.19476 1.42 0.158
## Year2000 0.37350 0.19146 1.95 0.052 .
## Year2001 0.11364 0.21584 0.53 0.599
## Year2002 0.07436 0.19653 0.38 0.705
## Year2003 -0.09430 0.21844 -0.43 0.666
## Year2004 -0.03142 0.20412 -0.15 0.878
## Year2005 0.06199 0.19167 0.32 0.747
```

```

## Year2006          0.14948      0.19117      0.78      0.435
## Year2007          0.20247      0.18882      1.07      0.284
## Year2008          0.20218      0.19517      1.04      0.301
## Year2009          0.33755      0.20759      1.63      0.105
## Year2010          0.25288      0.24066      1.05      0.294
## Year2011         -0.00582      0.25353     -0.02      0.982
## Year2012          0.33723      0.20232      1.67      0.097 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.103, Adjusted R-squared:  0.0535
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 319 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##   0.104  0.866  0.952  0.892  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      2.91e-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 1.277 1      1.130
## Year              1.277 16      1.008

```


Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3091 -0.2906 0.0182 0.3013 1.7492
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9618 0.1772 5.43 1.1e-07 ***
## FirstAuthorFemale1 0.1083 0.0754 1.44 0.152
## Year1997 0.2184 0.2151 1.02 0.311
## Year1998 0.4367 0.2636 1.66 0.099 .
## Year1999 0.2607 0.1937 1.35 0.179
## Year2000 0.3817 0.1912 2.00 0.047 *
## Year2001 0.1226 0.2162 0.57 0.571
## Year2002 0.0721 0.1979 0.36 0.716
## Year2003 -0.1013 0.2231 -0.45 0.650
## Year2004 -0.0242 0.2062 -0.12 0.907
## Year2005 0.0579 0.1923 0.30 0.763
## Year2006 0.1523 0.1912 0.80 0.426
```

```

## Year2007          0.2087      0.1897      1.10      0.272
## Year2008          0.1891      0.1959      0.97      0.335
## Year2009          0.3473      0.2083      1.67      0.096 .
## Year2010          0.2333      0.2422      0.96      0.336
## Year2011         -0.0053      0.2548     -0.02      0.983
## Year2012          0.3265      0.2019      1.62      0.107
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.095, Adjusted R-squared:  0.0478
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 321 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0945 0.8650 0.9520 0.8920 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      2.91e-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 1.347 1      1.161
## Year      1.347 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3539 -0.2749  0.0308  0.2960  1.7243

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.98670    0.17573   5.61 4.2e-08 ***
## LastAuthorFemale1 -0.04622    0.06827  -0.68  0.499
## Year1997        0.20858    0.21518   0.97  0.333
## Year1998        0.41710    0.26918   1.55  0.122
## Year1999        0.27242    0.19331   1.41  0.160
## Year2000        0.37010    0.18975   1.95  0.052 .
## Year2001        0.10912    0.21916   0.50  0.619
## Year2002        0.06608    0.19437   0.34  0.734
## Year2003       -0.09901    0.21956  -0.45  0.652
## Year2004       -0.01356    0.20690  -0.07  0.948
## Year2005        0.06692    0.19013   0.35  0.725
## Year2006        0.13582    0.18941   0.72  0.474
## Year2007        0.22396    0.18791   1.19  0.234
## Year2008        0.22221    0.19408   1.14  0.253
## Year2009        0.36724    0.20887   1.76  0.080 .
## Year2010        0.26995    0.23641   1.14  0.254
## Year2011        0.00561    0.24965   0.02  0.982
## Year2012        0.34243    0.20145   1.70  0.090 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.0885, Adjusted R-squared:  0.041
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 315 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.106  0.862  0.953  0.889  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      2.91e-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: 344"
## [1] ""

```

```

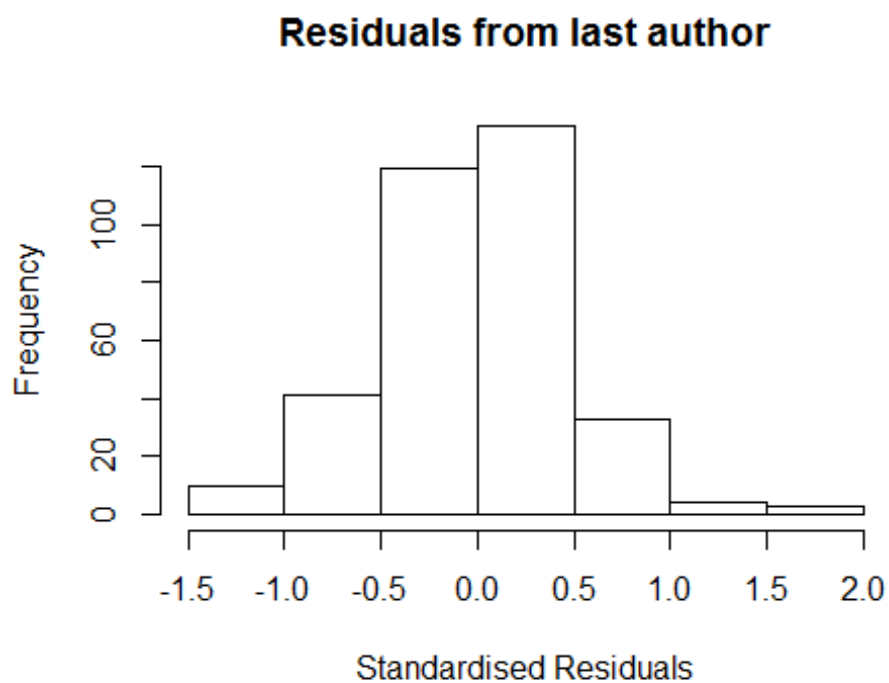
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2301"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    8    8    5    5    9   12    1    6    4   10   10    4   11
## 2011 2012
##   14   19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    3    5    3    1    8   10    1    6    2   10    9    4    7
## 2011 2012
##    7   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    3    5    3    1    8   10    1    6    2    8    9    3    7
## 2011 2012
##    6   10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.28942848510666"
## [1] "Male first author team size 2018 geometric mean: 2.35354689365025"

## 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 = 1
## 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: 2.37453746865839"

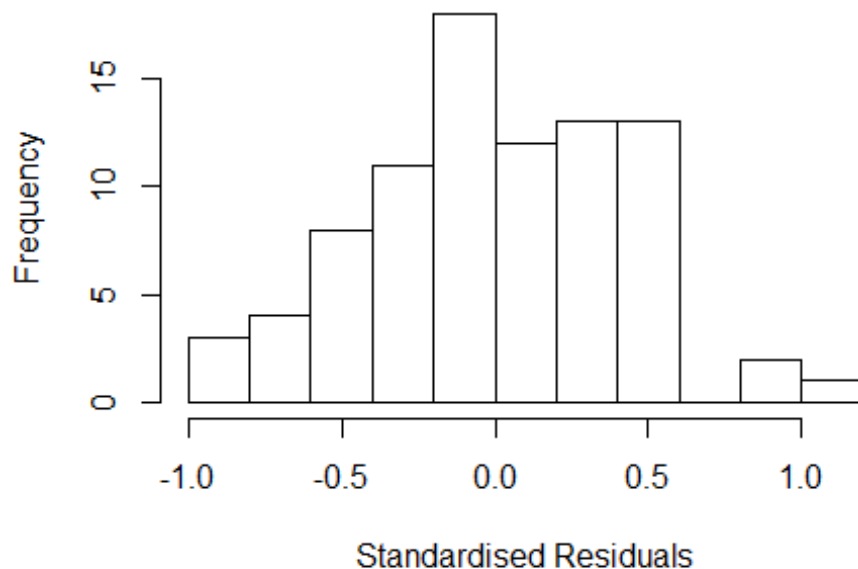
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3.5, 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 1.248e+13  1      3.533e+06
## LastAuthorFemale  2.188e+13  1      4.677e+06
## UniqueAuthors    1.229e+14  4      5.770e+01
## Year              4.601e+27 16      7.319e+00
```

Residuals from first and last author and team size



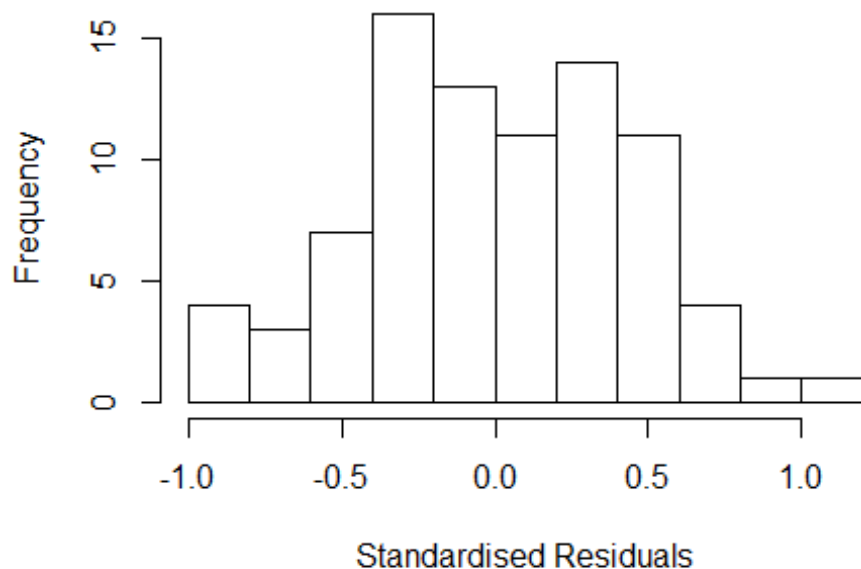
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.06e-01 -2.37e-01 3.33e-16 3.36e-01 1.09e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6798 0.1580 10.63 1.3e-15 ***
## FirstAuthorFemale1 -0.0306 0.1102 -0.28 0.78240
## LastAuthorFemale1 0.0840 0.1158 0.73 0.47069
## UniqueAuthors2 0.2678 0.1398 1.92 0.05994 .
## UniqueAuthors3 0.0407 0.1506 0.27 0.78766
## UniqueAuthors4 -0.2577 0.2774 -0.93 0.35642
## UniqueAuthors5 -0.0989 0.2279 -0.43 0.66580
## Year1997 -0.9983 0.1701 -5.87 1.9e-07 ***
## Year1998 -0.4832 0.4616 -1.05 0.29922
## Year1999 -0.5681 0.2747 -2.07 0.04284 *
```

```

## Year2000          -0.1725      0.2328    -0.74    0.46166
## Year2001          -0.1882      0.1924    -0.98    0.33178
## Year2002          -0.3739      0.1951    -1.92    0.05987 .
## Year2003          -0.8317      0.2144    -3.88    0.00026 ***
## Year2004          -1.3748      0.1580    -8.70    2.4e-12 ***
## Year2005          -0.4919      0.1630    -3.02    0.00369 **
## Year2006          -0.8905      0.3073    -2.90    0.00519 **
## Year2007          -0.8494      0.1858    -4.57    2.4e-05 ***
## Year2008          -0.4474      0.2077    -2.15    0.03511 *
## Year2009          -0.3097      0.2507    -1.24    0.22148
## Year2010          -0.4084      0.2484    -1.64    0.10523
## Year2011          -0.5001      0.2669    -1.87    0.06567 .
## Year2012          -0.3293      0.1987    -1.66    0.10244
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.522
## Multiple R-squared:  0.292, Adjusted R-squared:  0.0411
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 78 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.642  0.919   0.962   0.941   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          1.18e-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 -1.242e+17 1      NaN
## LastAuthorFemale  -1.616e+17 1      NaN
## Year               -5.708e+29 16      NaN

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -9.25e-01 -2.98e-01  2.22e-16  2.89e-01  1.09e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.9389     0.1347   14.40 < 2e-16 ***
## FirstAuthorFemale1 -0.0219     0.1347   -0.16  0.87125
## LastAuthorFemale1  0.1215     0.1285    0.95  0.34800
## Year1997          -1.2574     0.1487   -8.46  4.0e-12 ***
## Year1998           -0.7593     0.4348   -1.75  0.08541 .
## Year1999           -0.7580     0.2787   -2.72  0.00835 **
## Year2000           -0.4342     0.1903   -2.28  0.02572 *
## Year2001           -0.4935     0.1285   -3.84  0.00028 ***
## Year2002           -0.6084     0.2016   -3.02  0.00361 **
## Year2003           -1.0775     0.2145   -5.02  4.1e-06 ***
## Year2004           -1.6339     0.1347  -12.13 < 2e-16 ***
## Year2005           -0.6464     0.1782   -3.63  0.00056 ***
```



```

## Year2006          -1.1727      0.3042   -3.85  0.00026 ***
## Year2007          -1.0143      0.2185   -4.64  1.7e-05 ***
## Year2008          -0.6637      0.2239   -2.96  0.00421 **
## Year2009          -0.6322      0.2191   -2.89  0.00527 **
## Year2010          -0.6639      0.2190   -3.03  0.00347 **
## Year2011          -0.7764      0.1741   -4.46  3.3e-05 ***
## Year2012          -0.5691      0.1981   -2.87  0.00548 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.242, Adjusted R-squared:  0.0357
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.641  0.923  0.956  0.930  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      1.18e-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 -2.707e+24 1      NaN
## Year              -2.707e+24 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 = "MM"
## Residuals:
##      Min      1Q    Median      3Q      Max

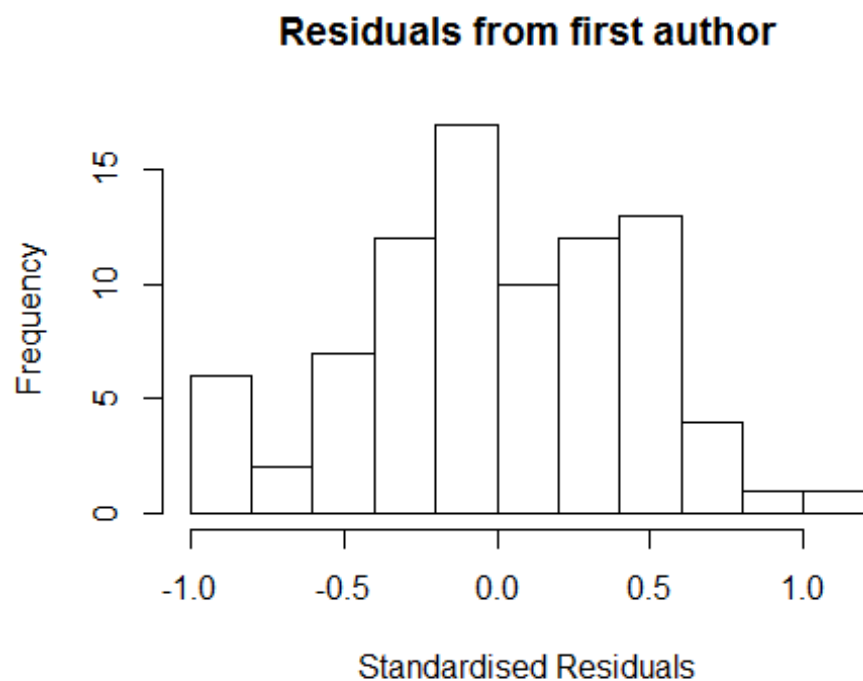
```

```

## -9.56e-01 -2.97e-01 -2.22e-15  3.12e-01  1.09e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.87e+00   1.11e-01  1.69e+01 < 2e-16 ***
## FirstAuthorFemale1 4.50e-02   1.11e-01  4.00e-01  0.68692
## Year1997        -1.19e+00   1.28e-01 -9.32e+00  1.0e-13 ***
## Year1998        -6.68e-01   4.42e-01 -1.51e+00  0.13524
## Year1999        -6.90e-01   2.69e-01 -2.57e+00  0.01247 *
## Year2000        -3.75e-01   2.05e-01 -1.83e+00  0.07147 .
## Year2001        -3.72e-01   3.16e-09 -1.18e+08 < 2e-16 ***
## Year2002        -5.60e-01   1.89e-01 -2.96e+00  0.00429 **
## Year2003        -1.00e+00   1.89e-01 -5.31e+00  1.3e-06 ***
## Year2004        -1.57e+00   1.11e-01 -1.41e+01 < 2e-16 ***
## Year2005        -5.49e-01   1.12e-01 -4.92e+00  6.0e-06 ***
## Year2006        -1.08e+00   2.63e-01 -4.10e+00  0.00011 ***
## Year2007        -9.16e-01   1.88e-01 -4.88e+00  6.8e-06 ***
## Year2008        -5.65e-01   1.78e-01 -3.17e+00  0.00227 **
## Year2009        -5.31e-01   2.01e-01 -2.64e+00  0.01029 *
## Year2010        -6.00e-01   2.10e-01 -2.86e+00  0.00570 **
## Year2011        -7.17e-01   1.65e-01 -4.34e+00  4.9e-05 ***
## Year2012        -4.68e-01   1.54e-01 -3.04e+00  0.00342 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.236, Adjusted R-squared:  0.0427
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.628  0.915  0.961   0.932   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           1.18e-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"

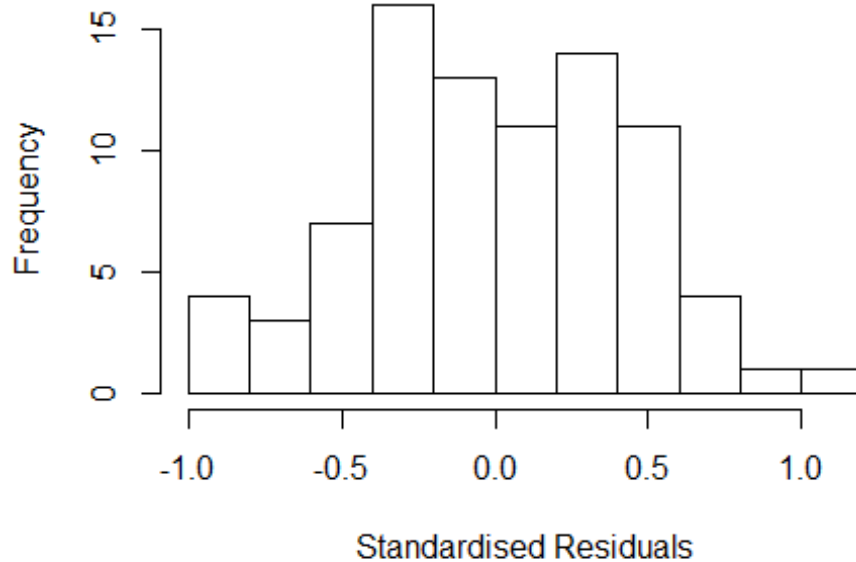
```

```
## 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 1.708e+13  1      4.133e+06
## Year             1.708e+13 16      2.591e+00
```

Residuals from last author



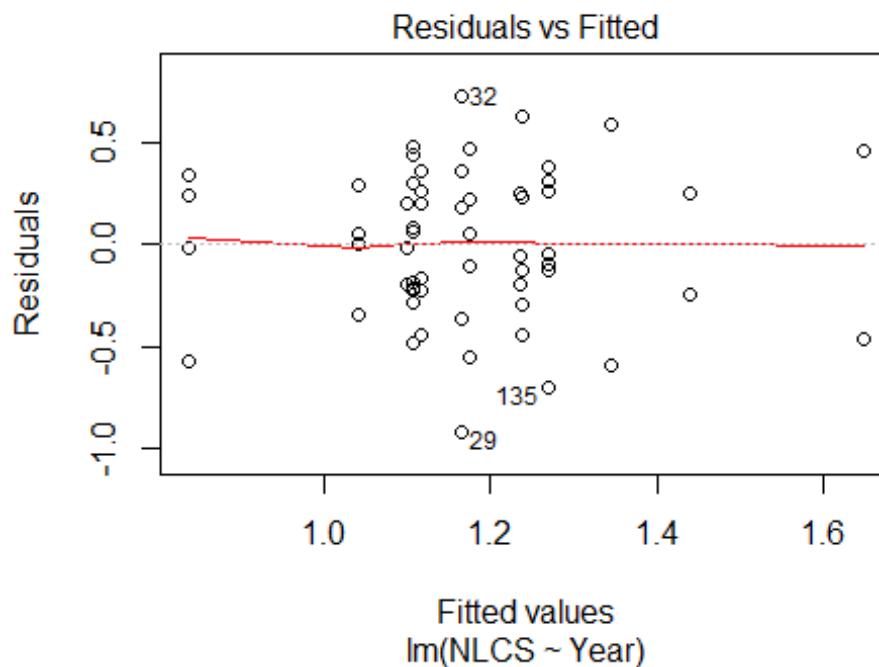
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.23e-01 -3.02e-01 1.33e-15 2.86e-01 1.09e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.92e+00 0.00e+00 Inf < 2e-16 ***
## LastAuthorFemale1 1.10e-01 1.05e-01 1.05e+00 0.29871
## Year1997 -1.24e+00 6.29e-02 -1.96e+01 < 2e-16 ***
## Year1998 -7.41e-01 4.19e-01 -1.77e+00 0.08153 .
## Year1999 -7.36e-01 2.42e-01 -3.04e+00 0.00340 **
## Year2000 -4.23e-01 1.83e-01 -2.31e+00 0.02397 *
## Year2001 -4.82e-01 1.05e-01 -4.60e+00 1.9e-05 ***
## Year2002 -5.99e-01 1.84e-01 -3.26e+00 0.00178 **
## Year2003 -1.06e+00 1.62e-01 -6.52e+00 1.1e-08 ***
## Year2004 -1.61e+00 3.93e-08 -4.10e+07 < 2e-16 ***
## Year2005 -6.29e-01 1.17e-01 -5.35e+00 1.1e-06 ***
## Year2006 -1.16e+00 2.91e-01 -3.98e+00 0.00017 ***
```

```

## Year2007          -9.94e-01    1.92e-01 -5.17e+00    2.3e-06 ***
## Year2008          -6.43e-01    1.88e-01 -3.41e+00    0.00109 **
## Year2009          -6.17e-01    1.91e-01 -3.24e+00    0.00189 **
## Year2010          -6.48e-01    2.10e-01 -3.08e+00    0.00298 **
## Year2011          -7.67e-01    1.81e-01 -4.23e+00    7.3e-05 ***
## Year2012          -5.48e-01    1.31e-01 -4.17e+00    8.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.518
## Multiple R-squared:  0.242, Adjusted R-squared:  0.05
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.637  0.922  0.955  0.931  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      1.18e-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: 85"
## [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
##    4    3    5    7    6    8    6    7    5    6    4   11    7    5   13
## 2011 2012
##   11    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    2    2    3    3    4    2    2    3    0    6    6    4   10
## 2011 2012

```

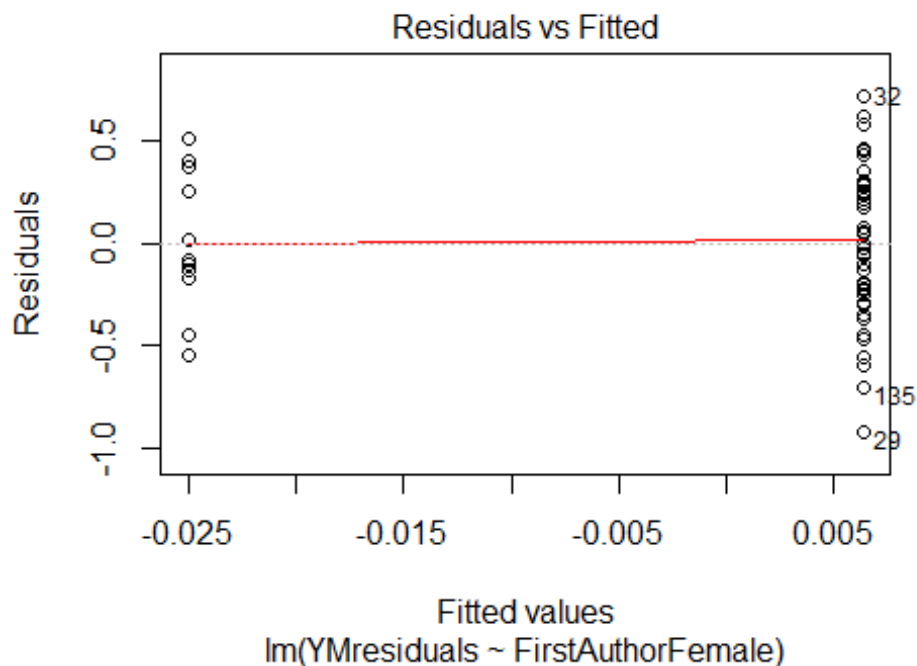
```
##      7      5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    1    2    3    1    4    1    1    1    0    6    6    3    10
## 2011 2012
##    6    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 9.3, df = 13, p-value = 0.8
```



```
##
## 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: 4.48140474655716"
## [1] "Male first author team size 2018 geometric mean: 2.82614631094813"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

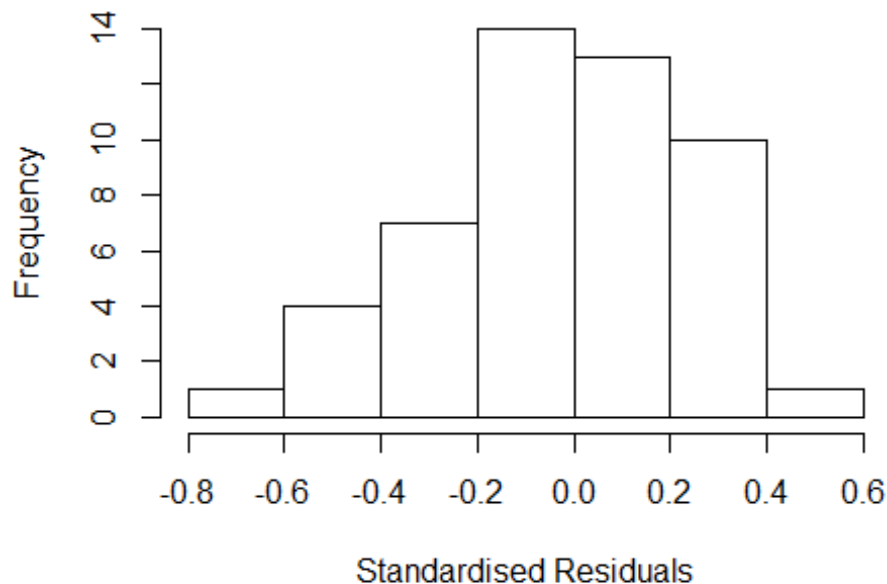
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.24264068711928"
## [1] "Male last author team size 2018 geometric mean: 3.03505497531656"

## 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 = 11, 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 -9.478e+14  1      NaN
## LastAuthorFemale -1.345e+15  1      NaN
## UniqueAuthors    -4.965e+30  4      NaN
## Year              -5.489e+43 13      NaN
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -6.05e-01 -1.79e-01 2.41e-15 1.66e-01 4.59e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.61143 0.15586 10.34 2.1e-11 ***
## FirstAuthorFemale1 0.01132 0.10976 0.10 0.9186
## LastAuthorFemale1 0.00627 0.13144 0.05 0.9623
## UniqueAuthors2 -0.05341 0.15057 -0.35 0.7253
## UniqueAuthors3 0.32457 0.15586 2.08 0.0459 *
## UniqueAuthors4 0.46536 0.15628 2.98 0.0057 **
## UniqueAuthors5 -0.02607 0.15946 -0.16 0.8712
## Year1999 0.08635 0.49983 0.17 0.8640
## Year2000 -0.49013 0.16604 -2.95 0.0061 **
## Year2001 -1.30902 0.12729 -10.28 2.4e-11 ***
```



```

## Year2002          -0.52829    0.18779    -2.81    0.0086 **
## Year2003          -0.80643    0.15586    -5.17    1.4e-05 ***
## Year2004          -0.36702    0.12729    -2.88    0.0072 **
## Year2005          -0.59202    0.20030    -2.96    0.0060 **
## Year2007          -0.69095    0.10708    -6.45    4.0e-07 ***
## Year2008          -0.58211    0.17112    -3.40    0.0019 **
## Year2009          -0.82611    0.24751    -3.34    0.0023 **
## Year2010          -0.69444    0.12961    -5.36    8.5e-06 ***
## Year2011          -0.57126    0.11384    -5.02    2.2e-05 ***
## Year2012          -0.55746    0.16166    -3.45    0.0017 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.539, Adjusted R-squared:  0.247
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.756  0.914  0.961   0.943  0.983   0.997
## Algorithmic parameters:
##           tuning.chi                bb           tuning.psi           refine.tol
##           1.55e+00                5.00e-01           4.69e+00           1.00e-07
##           rel.tol                solve.tol           eps.outlier           eps.x
##           1.00e-07                1.00e-07           2.00e-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 -1.123e+15  1         NaN
## LastAuthorFemale  -7.888e+14  1         NaN
## Year               -2.404e+15 13         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 =

```

```

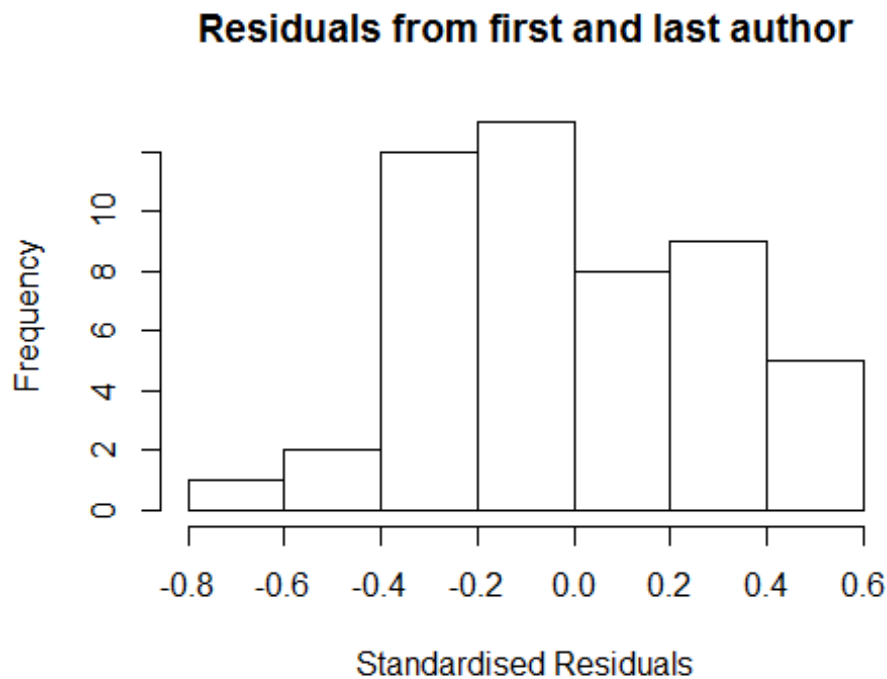
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q        Max
## -7.09e-01 -2.27e-01 -4.44e-16  2.43e-01  5.95e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.94e+00   3.27e-08  5.91e+07 < 2e-16 ***
## FirstAuthorFemale1 -7.47e-02   1.48e-01 -5.00e-01  0.6174
## LastAuthorFemale1  1.78e-01   1.24e-01  1.44e+00  0.1600
## Year1999        -3.78e-01   3.36e-01 -1.12e+00  0.2685
## Year2000        -8.11e-01   1.18e-01 -6.87e+00  6.5e-08 ***
## Year2001       -1.69e+00   1.71e-08 -9.85e+07 < 2e-16 ***
## Year2002       -9.41e-01   1.28e-01 -7.35e+00  1.6e-08 ***
## Year2003       -1.13e+00   2.95e-08 -3.83e+07 < 2e-16 ***
## Year2004       -7.45e-01   3.27e-08 -2.28e+07 < 2e-16 ***
## Year2005       -1.00e+00   1.57e-01 -6.37e+00  2.9e-07 ***
## Year2007       -7.60e-01   1.63e-01 -4.65e+00  4.9e-05 ***
## Year2008       -8.65e-01   1.67e-01 -5.17e+00  1.0e-05 ***
## Year2009       -1.22e+00   2.46e-01 -4.97e+00  1.9e-05 ***
## Year2010       -8.63e-01   1.11e-01 -7.79e+00  4.6e-09 ***
## Year2011       -6.54e-01   1.84e-01 -3.56e+00  0.0011 **
## Year2012       -7.65e-01   1.58e-01 -4.84e+00  2.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.378, Adjusted R-squared:  0.103
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 42 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.685  0.907  0.954  0.928  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      2.00e-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"

## 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 13                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
## -7.52e-01 -2.06e-01 -6.11e-16  2.46e-01  6.59e-01
##
```

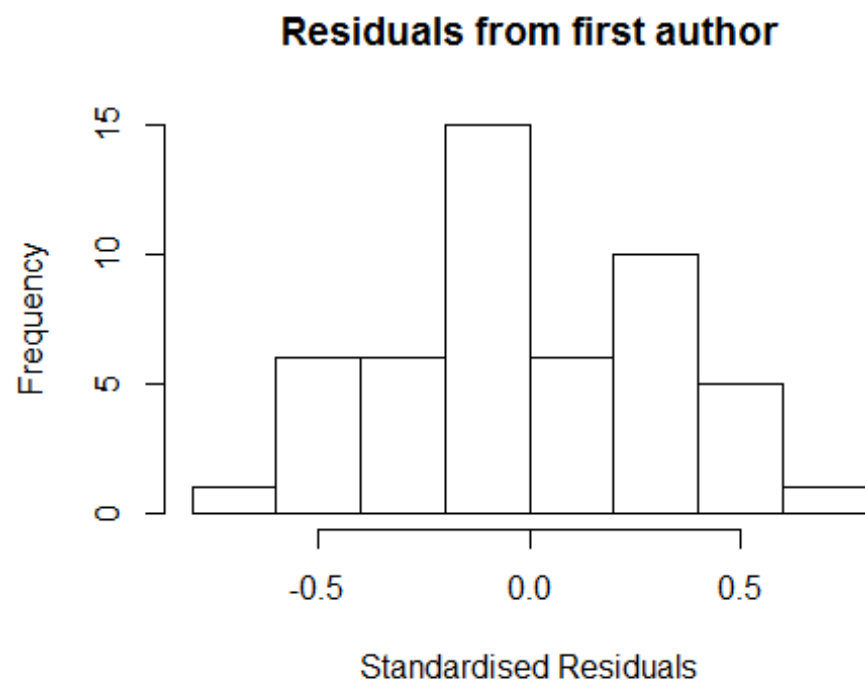
```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.93600    0.00000      Inf < 2e-16 ***
## FirstAuthorFemale1 -0.00852    0.13835    -0.06  0.9513
## Year1999          -0.28850    0.47905    -0.60  0.5509
## Year2000          -0.83399    0.11431    -7.30  1.6e-08 ***
## Year2001          -1.68700    0.00000    -Inf < 2e-16 ***
## Year2002          -0.89125    0.11717    -7.61  6.4e-09 ***
## Year2003          -1.13100    0.00000    -Inf < 2e-16 ***
## Year2004          -0.74500    0.00000    -Inf < 2e-16 ***
## Year2005          -0.89048    0.13835    -6.44  2.1e-07 ***
## Year2007          -0.75075    0.16100    -4.66  4.4e-05 ***
## Year2008          -0.81350    0.13736    -5.92  9.8e-07 ***
## Year2009          -1.16208    0.24362    -4.77  3.2e-05 ***
## Year2010          -0.83325    0.10453    -7.97  2.2e-09 ***
## Year2011          -0.61083    0.17451    -3.50  0.0013 **
## Year2012          -0.72239    0.20987    -3.44  0.0015 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.364
## Multiple R-squared:  0.35, Adjusted R-squared:  0.0902
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 43 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.649  0.877  0.956  0.926  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      2.00e-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"

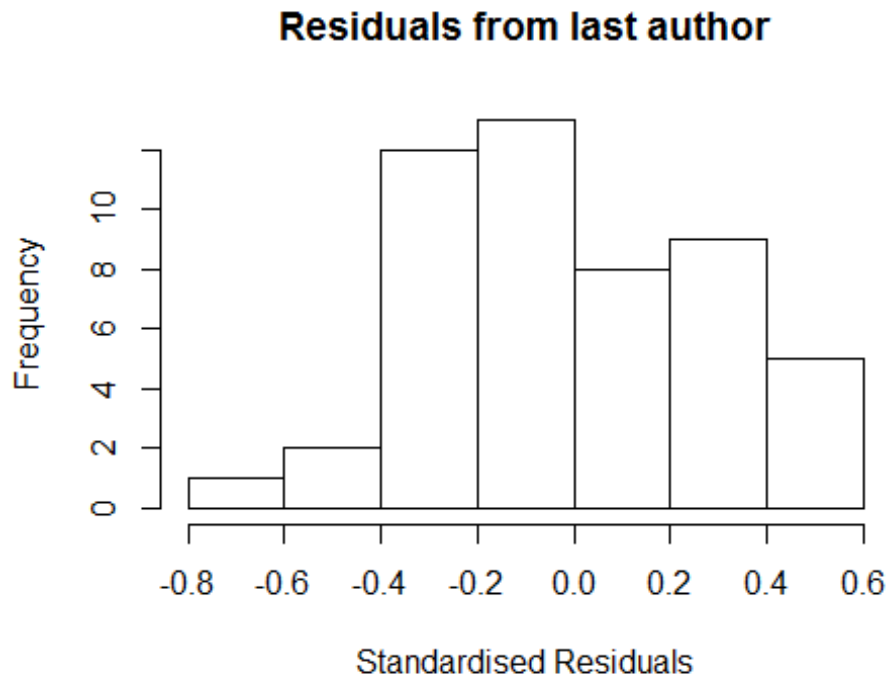
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

```

```
## 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 13          NaN
```



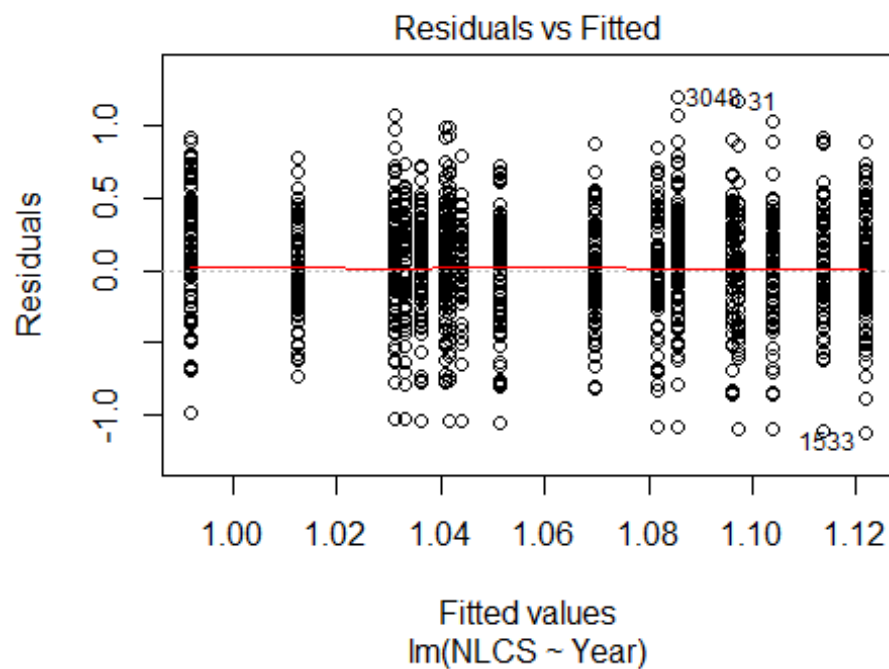
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -6.80e-01 -2.43e-01 -6.66e-16 2.51e-01 5.51e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.9360 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 0.1573 0.1093 1.44 0.15904
## Year1999 -0.3672 0.3268 -1.12 0.26888
## Year2000 -0.8369 0.0964 -8.68 3.0e-10 ***
## Year2001 -1.6870 0.0000 -Inf < 2e-16 ***
## Year2002 -0.9345 0.1240 -7.53 7.9e-09 ***
## Year2003 -1.1310 0.0000 -Inf < 2e-16 ***
## Year2004 -0.7450 0.0000 -Inf < 2e-16 ***
## Year2005 -1.0563 0.1093 -9.66 2.1e-11 ***
## Year2007 -0.7865 0.1392 -5.65 2.2e-06 ***
## Year2008 -0.8718 0.1590 -5.48 3.7e-06 ***
## Year2009 -1.2621 0.2531 -4.99 1.7e-05 ***
```

```

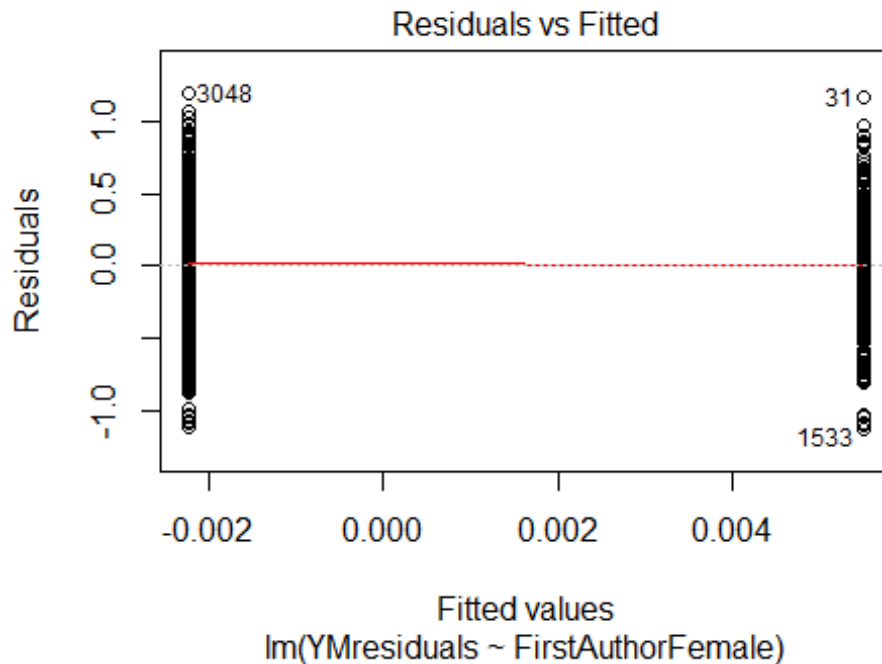
## Year2010          -0.8662      0.1098    -7.89  2.8e-09 ***
## Year2011          -0.6828      0.1690    -4.04  0.00028 ***
## Year2012          -0.7711      0.1496    -5.15  1.0e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.365, Adjusted R-squared:  0.111
## Convergence in 9 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.778  0.926  0.963   0.946  0.982   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      2.00e-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: 50"
## [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
## 129 165 131 178 151 198 171 173 138 135 154 170 141 147 155
## 2011 2012
## 175 194
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 59 79 67 92 62 75 103 112 90 92 117 112 97 108 107
## 2011 2012
## 114 129
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```
## 52 72 64 81 53 72 94 102 85 76 103 95 82 96 98
## 2011 2012
## 106 122
## [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
```

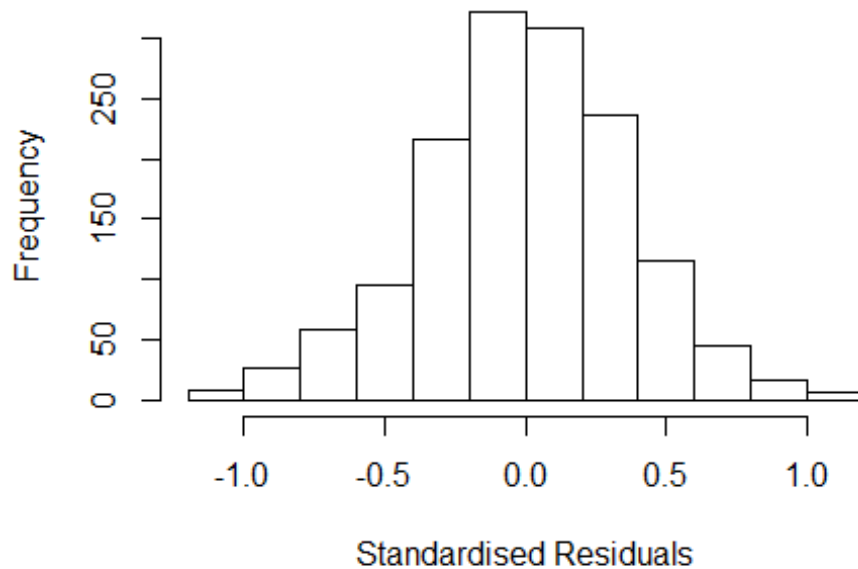


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 4.4, df = 1, p-value = 0.04
```

```
## [1] "Female first author team size 2018 geometric mean: 3.37279809711066"
## [1] "Male first author team size 2018 geometric mean: 3.36952757068956"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1100, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.20968052520874"
## [1] "Male last author team size 2018 geometric mean: 3.41950072421596"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 880, 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 1.167 1      1.080
## LastAuthorFemale  1.138 1      1.067
## UniqueAuthors    1.488 4      1.051
## Year              1.799 16     1.019
```

Residuals from first and last author and team size



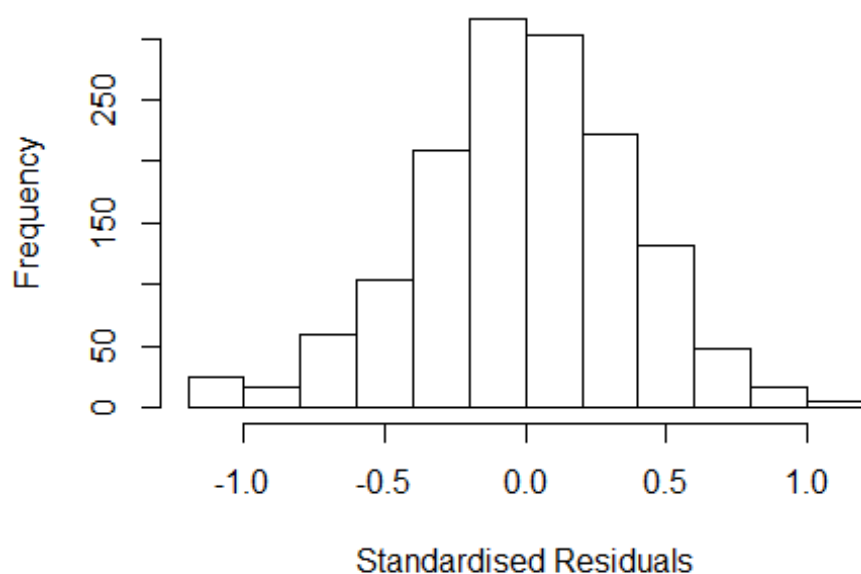
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.15826 -0.22627 0.00361 0.23710 1.18644
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9773 0.0620 15.75 < 2e-16 ***
## FirstAuthorFemale1 0.0169 0.0216 0.78 0.43561
## LastAuthorFemale1 -0.0574 0.0238 -2.41 0.01588 *
## UniqueAuthors2 0.1262 0.0319 3.95 8.1e-05 ***
## UniqueAuthors3 0.1254 0.0334 3.76 0.00018 ***
## UniqueAuthors4 0.1908 0.0366 5.22 2.1e-07 ***
## UniqueAuthors5 0.2889 0.0400 7.23 8.1e-13 ***
## Year1997 0.0550 0.0747 0.74 0.46166
## Year1998 -0.0372 0.0780 -0.48 0.63355
## Year1999 0.0157 0.0737 0.21 0.83163
```

```

## Year2000          0.0066      0.0756      0.09  0.93037
## Year2001          0.0133      0.0759      0.18  0.86103
## Year2002          0.0268      0.0681      0.39  0.69431
## Year2003          0.0380      0.0697      0.55  0.58575
## Year2004         -0.0435      0.0711     -0.61  0.54061
## Year2005         -0.0621      0.0707     -0.88  0.38009
## Year2006         -0.0797      0.0717     -1.11  0.26634
## Year2007         -0.0345      0.0732     -0.47  0.63719
## Year2008         -0.1001      0.0722     -1.39  0.16567
## Year2009         -0.0536      0.0691     -0.78  0.43822
## Year2010         -0.0788      0.0747     -1.05  0.29197
## Year2011         -0.0187      0.0760     -0.25  0.80516
## Year2012         -0.1060      0.0774     -1.37  0.17121
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.0607, Adjusted R-squared:  0.0462
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 126 weights are ~= 1. The remaining 1327 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.217  0.867  0.952   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      6.88e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.155 1      1.075
## LastAuthorFemale  1.125 1      1.061
## Year              1.259 16      1.007

```

Residuals from first and last author



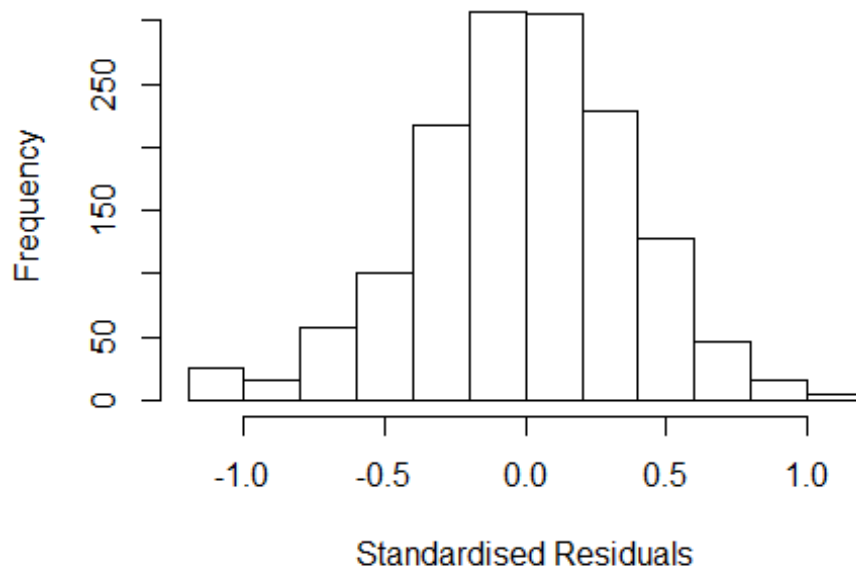
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.150091 -0.230005 -0.000701  0.237760  1.173941
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.07476    0.05952   18.06  <2e-16 ***
## FirstAuthorFemale1  0.02030    0.02188    0.93   0.354
## LastAuthorFemale1 -0.04837    0.02418   -2.00   0.046 *
## Year1997         0.06228    0.07393    0.84   0.400
## Year1998        -0.03647    0.07703   -0.47   0.636
## Year1999         0.01700    0.07418    0.23   0.819
## Year2000         0.00232    0.07551    0.03   0.976
## Year2001         0.02585    0.07681    0.34   0.737
## Year2002         0.04654    0.06890    0.68   0.499
## Year2003         0.05503    0.06910    0.80   0.426
## Year2004        -0.00885    0.07075   -0.13   0.900
## Year2005        -0.02431    0.06933   -0.35   0.726
```

```

## Year2006      -0.03452    0.07158   -0.48    0.630
## Year2007      0.00526    0.07310    0.07    0.943
## Year2008     -0.04340    0.07195   -0.60    0.546
## Year2009     -0.02325    0.06761   -0.34    0.731
## Year2010     -0.01906    0.07358   -0.26    0.796
## Year2011      0.05511    0.07432    0.74    0.459
## Year2012     -0.04393    0.07615   -0.58    0.564
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.356
## Multiple R-squared:  0.0124, Adjusted R-squared:  -8.66e-06
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 122 weights are ~= 1. The remaining 1331 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.253  0.867  0.952  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      6.88e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01      5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000      200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi      subsampling      cov
##      "bisquare"      "nonsingular"      ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.149 1      1.072
## Year      1.149 16      1.004

```

Residuals from first author



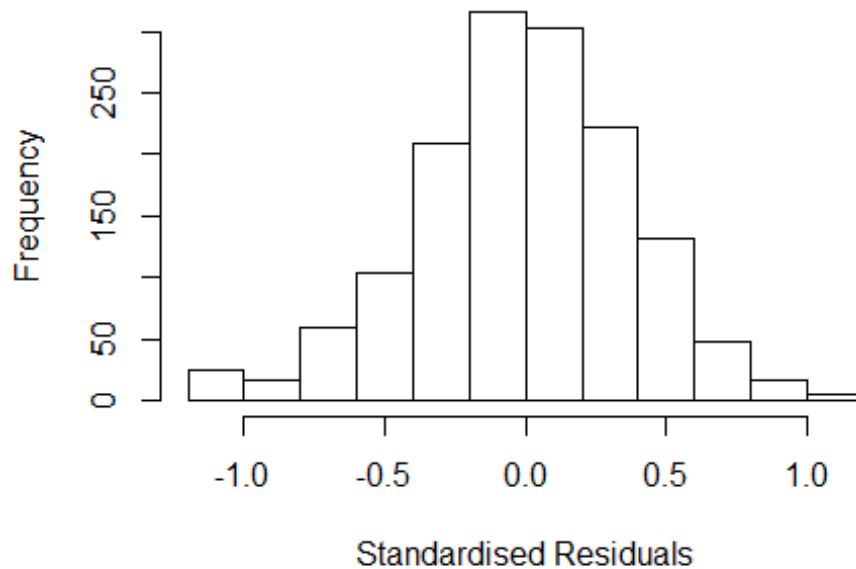
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.13581 -0.23333 0.00118 0.24168 1.18810
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06761 0.05985 17.84 <2e-16 ***
## FirstAuthorFemale1 0.01329 0.02196 0.61 0.55
## Year1997 0.05491 0.07387 0.74 0.46
## Year1998 -0.03928 0.07751 -0.51 0.61
## Year1999 0.01267 0.07454 0.17 0.87
## Year2000 0.00160 0.07597 0.02 0.98
## Year2001 0.02925 0.07749 0.38 0.71
## Year2002 0.04721 0.06946 0.68 0.50
## Year2003 0.05301 0.06953 0.76 0.45
## Year2004 -0.01213 0.07101 -0.17 0.86
## Year2005 -0.02373 0.06981 -0.34 0.73
## Year2006 -0.03129 0.07202 -0.43 0.66
```

```

## Year2007          0.00382    0.07357    0.05    0.96
## Year2008          -0.04440    0.07251   -0.61    0.54
## Year2009          -0.02991    0.06795   -0.44    0.66
## Year2010          -0.01880    0.07401   -0.25    0.80
## Year2011           0.04985    0.07499    0.66    0.51
## Year2012          -0.04544    0.07641   -0.59    0.55
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.00938,    Adjusted R-squared:  -0.00236
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 118 weights are ~= 1. The remaining 1335 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.248  0.871  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      6.88e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Regression 4: Last author gender, Year as factors"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.113 1          1.055
## Year            1.113 16          1.003

```

Residuals from last author



```
## [1] "List of 0 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
## -1.13812 -0.23222 0.00186 0.23878 1.19298
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.07602 0.05962 18.05 <2e-16 ***
## LastAuthorFemale1 -0.04484 0.02402 -1.87 0.062 .
## Year1997 0.06512 0.07358 0.89 0.376
## Year1998 -0.03495 0.07674 -0.46 0.649
## Year1999 0.01920 0.07391 0.26 0.795
## Year2000 0.00267 0.07532 0.04 0.972
## Year2001 0.02881 0.07654 0.38 0.707
## Year2002 0.04998 0.06845 0.73 0.465
## Year2003 0.05732 0.06878 0.83 0.405
## Year2004 -0.00289 0.06987 -0.04 0.967
## Year2005 -0.02181 0.06908 -0.32 0.752
## Year2006 -0.02993 0.07118 -0.42 0.674
```

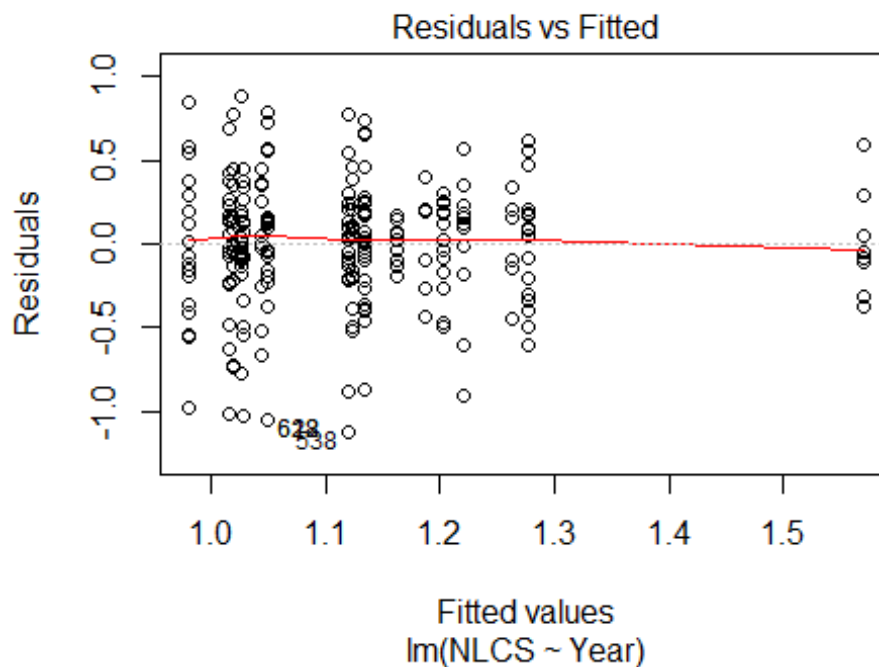


```

## Year2007      0.00900    0.07264    0.12    0.901
## Year2008     -0.03843    0.07127   -0.54    0.590
## Year2009     -0.01891    0.06704   -0.28    0.778
## Year2010     -0.01145    0.07226   -0.16    0.874
## Year2011      0.06210    0.07323    0.85    0.397
## Year2012     -0.03693    0.07467   -0.49    0.621
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.0119, Adjusted R-squared:  0.000153
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 127 weights are ~= 1. The remaining 1326 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.236  0.867  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      6.88e-05      1.82e-12
## warn.limit.reject warn.limit.meanrw
##      5.00e-01          5.00e-01
## nResample      max.it      best.r.s      k.fast.s      k.max maxit.scale
##      500          50          2          1          1000          200
## trace.lev      mts      compute.rd
##      0          1000          0
##      psi          subsampling          cov
##      "bisquare"          "nonsingular"          ".vcov.avar1"
## compute.outlier.stats
##      "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 1453"
## [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
##   28   28   28   46   37   39   30   39   33   52   30   31   37   34   33
## 2011 2012
##   27   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9   11    8    6   10    6   15   19   16   32   15   17   24   20   19
## 2011 2012

```

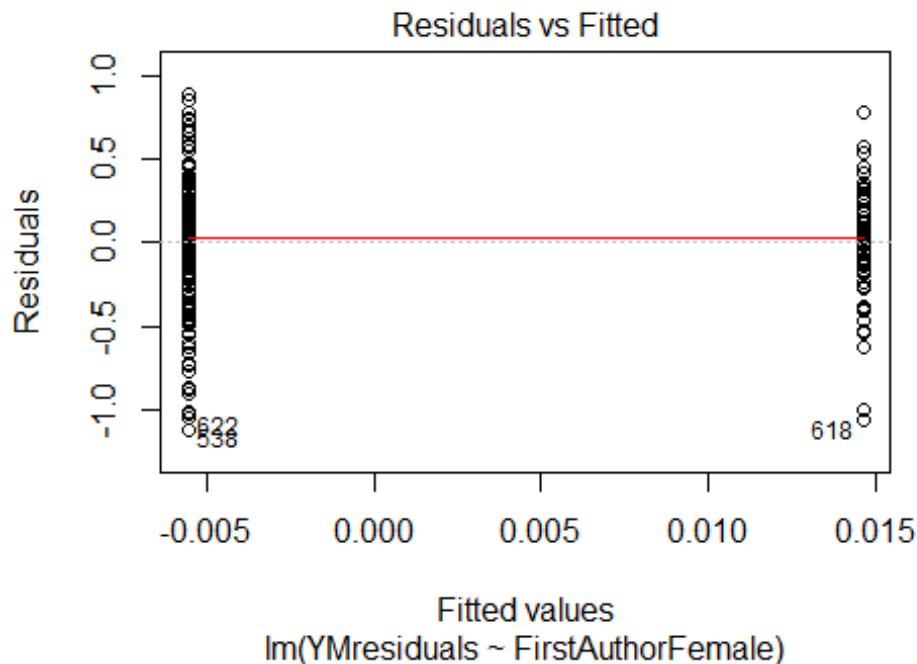
```
## 16 15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 11 6 6 9 4 14 16 16 31 13 15 21 19 18
## 2011 2012
## 14 15
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 22, df = 16, p-value = 0.1
```



```
##
## 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: 3.12977122905613"
## [1] "Male first author team size 2018 geometric mean: 2.52071780228707"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

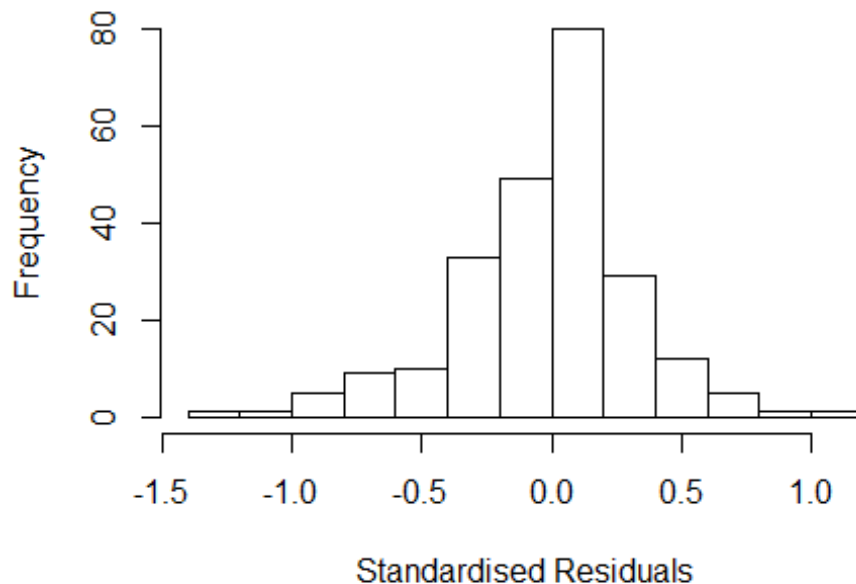
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.35215804504935"
## [1] "Male last author team size 2018 geometric mean: 2.92997628046227"

## 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 = 36, 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 1.593 1      1.262
## LastAuthorFemale  1.407 1      1.186
## UniqueAuthors    5.666 4      1.242
## Year             10.020 16     1.075
```

Residuals from first and last author and team size



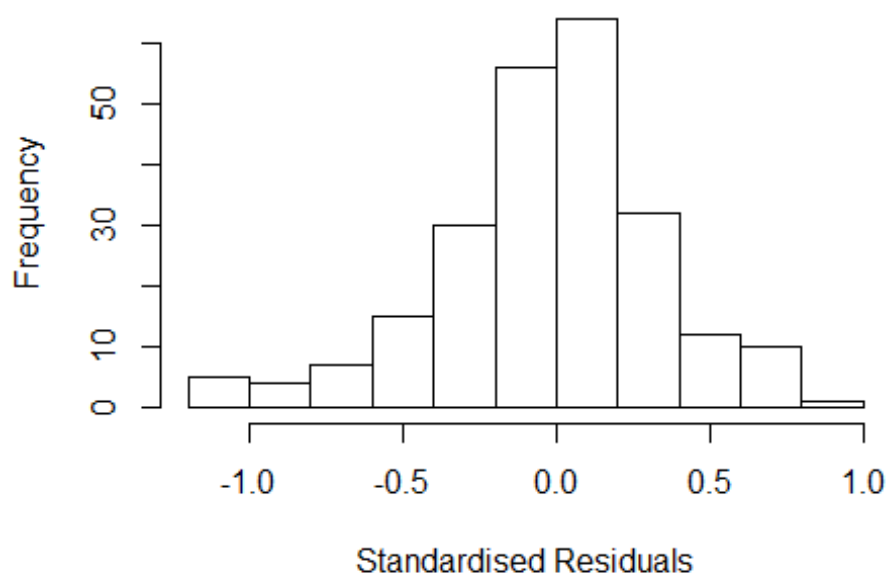
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2229 -0.1960  0.0196  0.1539  1.0816
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.95236    0.06157   15.47 < 2e-16 ***
## FirstAuthorFemale1 -0.00162    0.04645   -0.03  0.97220
## LastAuthorFemale1  0.02811    0.05057    0.56  0.57882
## UniqueAuthors2     0.18122    0.06639    2.73  0.00687 **
## UniqueAuthors3     0.26697    0.08016    3.33  0.00102 **
## UniqueAuthors4     0.10881    0.09712    1.12  0.26380
## UniqueAuthors5     0.23928    0.07086    3.38  0.00087 ***
## Year1997          0.14641    0.13556    1.08  0.28134
## Year1998          0.39922    0.11261    3.55  0.00048 ***
## Year1999          0.15049    0.08504    1.77  0.07822 .
```

```

## Year2000      -0.05226    0.16609   -0.31  0.75335
## Year2001      0.09507    0.22285    0.43  0.67011
## Year2002      0.13708    0.06408    2.14  0.03357 *
## Year2003     -0.21296    0.09465   -2.25  0.02546 *
## Year2004     -0.05474    0.10077   -0.54  0.58755
## Year2005     -0.00277    0.07181   -0.04  0.96925
## Year2006     -0.13041    0.06286   -2.07  0.03922 *
## Year2007     -0.04019    0.08064   -0.50  0.61876
## Year2008      0.08935    0.06535    1.37  0.17302
## Year2009     -0.05523    0.08557   -0.65  0.51933
## Year2010     -0.03503    0.13945   -0.25  0.80191
## Year2011      0.07712    0.11673    0.66  0.50955
## Year2012     -0.10311    0.09244   -1.12  0.26595
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.274
## Multiple R-squared:  0.216, Adjusted R-squared:  0.136
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 211 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0091 0.8520 0.9490 0.8620 0.9860 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          4.24e-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 1.447 1 1.203
## LastAuthorFemale 1.289 1 1.135
## Year 1.819 16 1.019

```

Residuals from first and last author



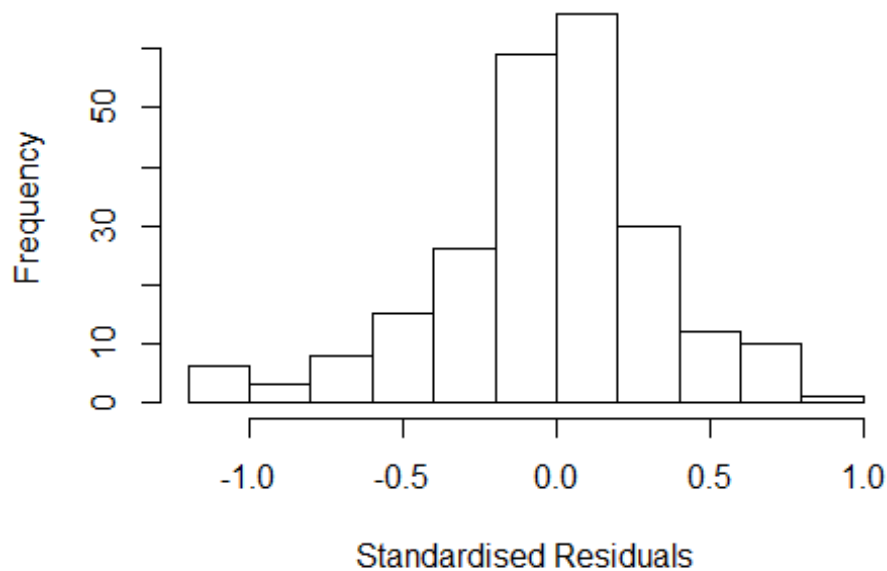
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.18918 -0.20929  0.00346  0.17818  0.95525
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1345     0.0402   28.20  <2e-16 ***
## FirstAuthorFemale1  0.0110     0.0476    0.23   0.8167
## LastAuthorFemale1  0.0223     0.0522    0.43   0.6691
## Year1997           0.1725     0.1120    1.54   0.1249
## Year1998           0.3424     0.1069    3.20   0.0016 **
## Year1999           0.1288     0.1228    1.05   0.2955
## Year2000          -0.0939     0.1548   -0.61   0.5446
## Year2001           0.0278     0.2278    0.12   0.9030
## Year2002           0.1088     0.0692    1.57   0.1174
## Year2003          -0.2688     0.1192   -2.25   0.0252 *
## Year2004          -0.0270     0.1027   -0.26   0.7930
## Year2005          -0.0343     0.0783   -0.44   0.6619
```

```

## Year2006          -0.1559      0.0644    -2.42    0.0164 *
## Year2007          -0.0190      0.0848    -0.22    0.8232
## Year2008           0.0547      0.0687     0.80    0.4267
## Year2009          -0.0853      0.0908    -0.94    0.3486
## Year2010          -0.0342      0.1152    -0.30    0.7669
## Year2011           0.1229      0.1120     1.10    0.2737
## Year2012          -0.1033      0.1006    -1.03    0.3061
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.291
## Multiple R-squared:  0.133, Adjusted R-squared:  0.0608
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 212 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0578 0.8240 0.9510 0.8670 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      4.24e-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 1.422 1      1.192
## Year              1.422 16      1.011

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.1934 -0.1995 0.0035 0.1742 0.9531
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1399 0.0398 28.67 <2e-16 ***
## FirstAuthorFemale1 0.0121 0.0479 0.25 0.8010
## Year1997 0.1721 0.1125 1.53 0.1275
## Year1998 0.3404 0.1055 3.22 0.0015 **
## Year1999 0.1295 0.1226 1.06 0.2923
## Year2000 -0.0989 0.1547 -0.64 0.5232
## Year2001 0.0223 0.2289 0.10 0.9224
## Year2002 0.1071 0.0706 1.52 0.1305
## Year2003 -0.2721 0.1206 -2.26 0.0251 *
## Year2004 -0.0293 0.1026 -0.29 0.7754
## Year2005 -0.0330 0.0790 -0.42 0.6768
## Year2006 -0.1610 0.0641 -2.51 0.0128 *
```

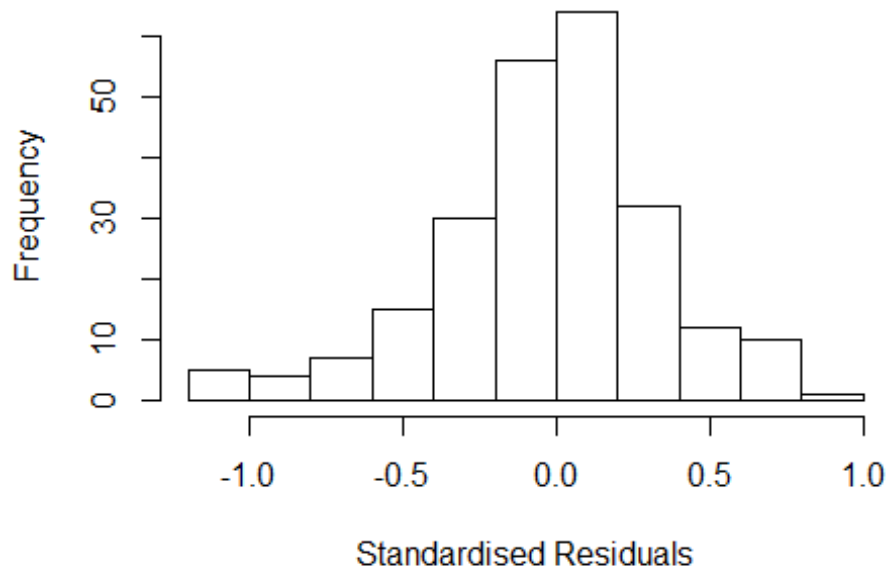


```

## Year2007          -0.0208      0.0857   -0.24   0.8081
## Year2008           0.0535      0.0692    0.77   0.4402
## Year2009          -0.0827      0.0914   -0.90   0.3665
## Year2010          -0.0354      0.1158   -0.31   0.7602
## Year2011           0.1229      0.1117    1.10   0.2724
## Year2012          -0.1073      0.1000   -1.07   0.2846
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.289
## Multiple R-squared:  0.134, Adjusted R-squared:  0.0665
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 211 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0489 0.8170 0.9510 0.8640 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      4.24e-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 1.272 1          1.128
## Year            1.272 16          1.008

```

Residuals from last author



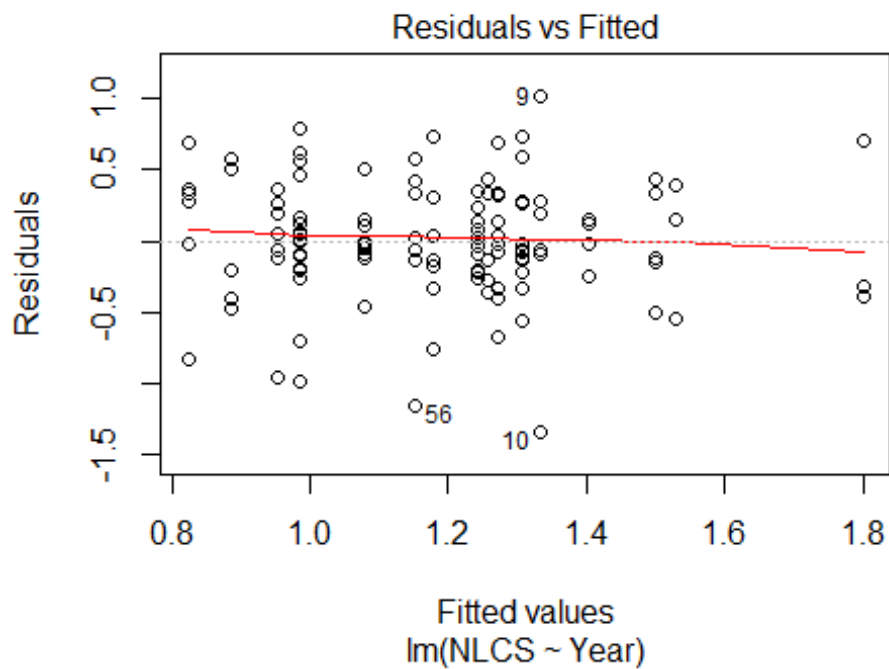
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.19305 -0.21226 0.00237 0.17821 0.95389
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1358 0.0393 28.86 <2e-16 ***
## LastAuthorFemale1 0.0229 0.0527 0.43 0.6646
## Year1997 0.1721 0.1114 1.54 0.1238
## Year1998 0.3426 0.1059 3.23 0.0014 **
## Year1999 0.1298 0.1235 1.05 0.2944
## Year2000 -0.0924 0.1532 -0.60 0.5472
## Year2001 0.0265 0.2284 0.12 0.9079
## Year2002 0.1090 0.0691 1.58 0.1162
## Year2003 -0.2687 0.1190 -2.26 0.0249 *
## Year2004 -0.0237 0.1027 -0.23 0.8174
## Year2005 -0.0320 0.0763 -0.42 0.6755
## Year2006 -0.1549 0.0642 -2.41 0.0167 *
```

```

## Year2007          -0.0168      0.0848   -0.20   0.8436
## Year2008           0.0573      0.0671    0.85   0.3946
## Year2009          -0.0818      0.0888   -0.92   0.3581
## Year2010          -0.0294      0.1131   -0.26   0.7949
## Year2011           0.1272      0.1108    1.15   0.2520
## Year2012          -0.0999      0.0971   -1.03   0.3047
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.29
## Multiple R-squared:  0.133, Adjusted R-squared:  0.0655
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 211 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0517 0.8190 0.9500 0.8650 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      4.24e-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: 236"
## [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
##    9    9   12   16   16   10   10   17   13   13   18   16   17   17   21
## 2011 2012
##   35   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    3    3    5    7    5    2    4    5   10    8    9    8    7    8
## 2011 2012

```

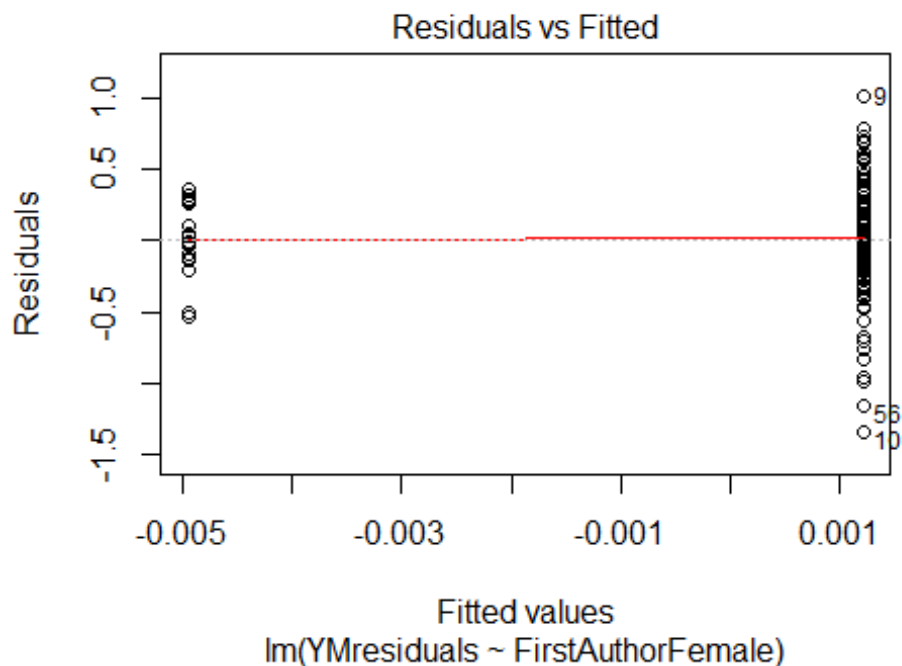
```
## 13 19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 3 2 4 7 5 2 2 5 7 6 8 8 6 8
## 2011 2012
## 12 17
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



```
##
## 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: 3.46410161513775"
## [1] "Male first author team size 2018 geometric mean: 2.7540002503904"
## 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 = 55, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.88449914061482"
## [1] "Male last author team size 2018 geometric mean: 3.06201519837311"

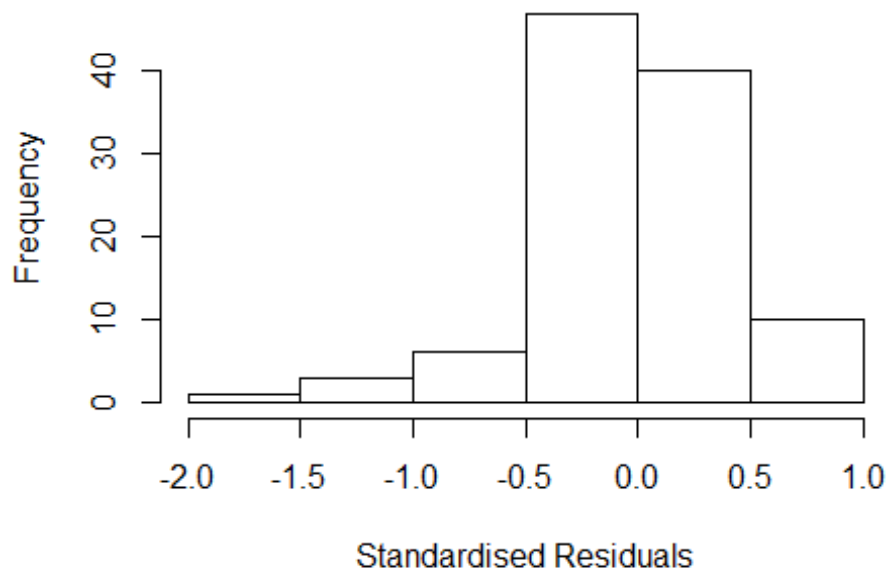
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 21, 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	3.611	1	1.900
## LastAuthorFemale	4.712	1	2.171
## UniqueAuthors	39.949	4	1.586
## Year	251.290	16	1.189

Residuals from first and last author and team size



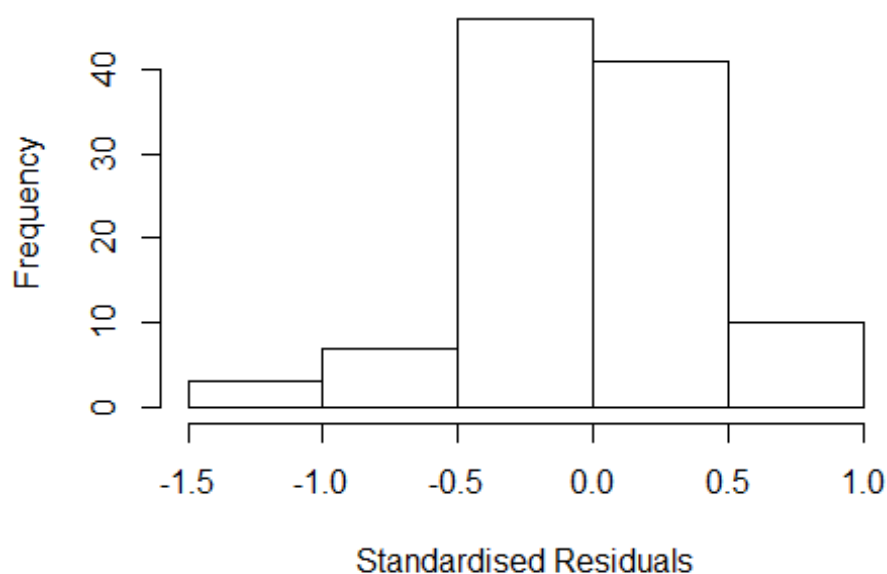
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5073 -0.2344 -0.0326 0.1824 0.8654
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4796 0.2935 5.04 2.6e-06 ***
## FirstAuthorFemale1 -0.0808 0.0928 -0.87 0.39
## LastAuthorFemale1 0.0823 0.1101 0.75 0.46
## UniqueAuthors2 0.0277 0.1866 0.15 0.88
## UniqueAuthors3 -0.0883 0.1616 -0.55 0.59
## UniqueAuthors4 -0.0280 0.1404 -0.20 0.84
## UniqueAuthors5 0.0492 0.1377 0.36 0.72
## Year1997 0.2500 0.4789 0.52 0.60
## Year1998 -0.0455 0.4330 -0.11 0.92
## Year1999 -0.0867 0.3321 -0.26 0.79
```

```

## Year2000          -0.1908      0.3591   -0.53      0.60
## Year2001          -0.2316      0.3399   -0.68      0.50
## Year2002          -0.2219      0.4004   -0.55      0.58
## Year2003          -0.1512      0.3273   -0.46      0.65
## Year2004          -0.6123      0.4144   -1.48      0.14
## Year2005          -0.2604      0.3133   -0.83      0.41
## Year2006          -0.4414      0.3300   -1.34      0.18
## Year2007          -0.4713      0.3206   -1.47      0.15
## Year2008          -0.3783      0.3669   -1.03      0.31
## Year2009          -0.5933      0.5954   -1.00      0.32
## Year2010          -0.1844      0.3511   -0.53      0.60
## Year2011          -0.1969      0.3438   -0.57      0.57
## Year2012          -0.4909      0.3200   -1.53      0.13
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.359
## Multiple R-squared:  0.23,   Adjusted R-squared:  0.0281
## Convergence in 28 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 99 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0393 0.8780 0.9610 0.8880 0.9920 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      9.35e-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 2.753 1      1.659
## LastAuthorFemale 3.366 1      1.835
## Year      8.358 16      1.069

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.483 -0.210 -0.018 0.229 0.862
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4826 0.2944 5.04 2.5e-06 ***
## FirstAuthorFemale1 -0.0714 0.0913 -0.78 0.44
## LastAuthorFemale1 0.0827 0.1133 0.73 0.47
## Year1997 0.2170 0.4810 0.45 0.65
## Year1998 -0.1114 0.3913 -0.28 0.78
## Year1999 -0.0978 0.3377 -0.29 0.77
## Year2000 -0.1855 0.3310 -0.56 0.58
## Year2001 -0.2369 0.3382 -0.70 0.49
## Year2002 -0.2111 0.4109 -0.51 0.61
## Year2003 -0.1784 0.3186 -0.56 0.58
## Year2004 -0.6264 0.4071 -1.54 0.13
## Year2005 -0.2738 0.3101 -0.88 0.38
```

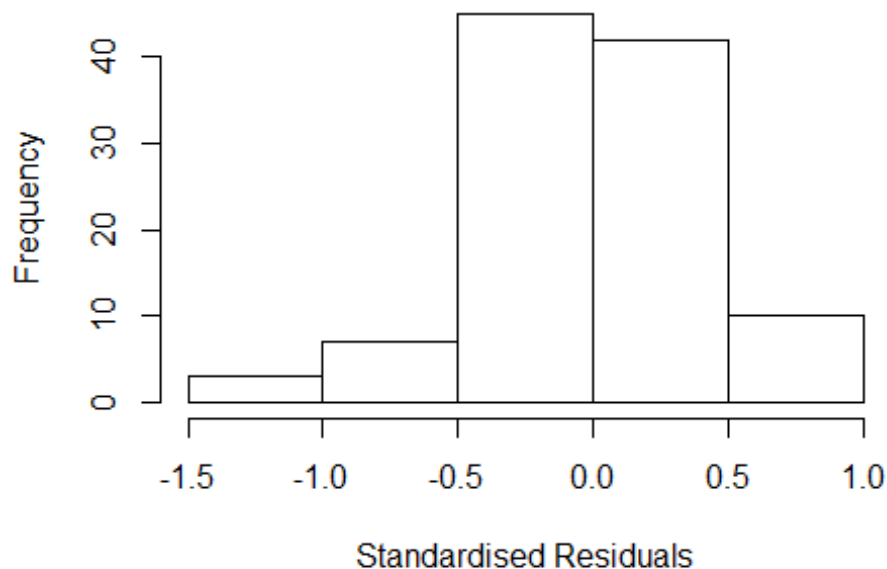


```

## Year2006          -0.4662      0.3249   -1.43      0.15
## Year2007          -0.4861      0.3077   -1.58      0.12
## Year2008          -0.4233      0.3288   -1.29      0.20
## Year2009          -0.6093      0.5608   -1.09      0.28
## Year2010          -0.2029      0.3431   -0.59      0.56
## Year2011          -0.2003      0.3193   -0.63      0.53
## Year2012          -0.4907      0.3110   -1.58      0.12
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.217, Adjusted R-squared:  0.0566
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 95 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0536 0.8660 0.9530 0.8830 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      9.35e-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 3.206 1      1.790
## Year              3.206 16      1.037

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.4881 -0.2225 -0.0101 0.2066 0.8569
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4881 0.2802 5.31 8e-07 ***
## FirstAuthorFemale1 -0.0554 0.0967 -0.57 0.57
## Year1997 0.1998 0.4695 0.43 0.67
## Year1998 -0.1249 0.3936 -0.32 0.75
## Year1999 -0.0895 0.3187 -0.28 0.78
## Year2000 -0.1688 0.3191 -0.53 0.60
## Year2001 -0.2264 0.3293 -0.69 0.49
## Year2002 -0.2166 0.3970 -0.55 0.59
## Year2003 -0.1919 0.3052 -0.63 0.53
## Year2004 -0.6351 0.3983 -1.59 0.11
## Year2005 -0.2401 0.2937 -0.82 0.42
## Year2006 -0.4709 0.3127 -1.51 0.14
```

```

## Year2007          -0.4520      0.2896   -1.56      0.12
## Year2008          -0.4206      0.3158   -1.33      0.19
## Year2009          -0.5443      0.5487   -0.99      0.32
## Year2010          -0.1865      0.3314   -0.56      0.58
## Year2011          -0.2008      0.3034   -0.66      0.51
## Year2012          -0.4916      0.2987   -1.65      0.10
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.215, Adjusted R-squared:  0.0647
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 92 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0295 0.8600 0.9480 0.8710 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      9.35e-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 3.927 1      1.982
## Year      3.927 16      1.044
##
## [1] "List of 0 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
## -1.4881 -0.2095 -0.0145  0.1982  0.8569

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4881    0.2800    5.31 7.8e-07 ***
## LastAuthorFemale1  0.0555    0.1218    0.46  0.650
## Year1997          0.1997    0.4704    0.42  0.672
## Year1998         -0.1526    0.4159   -0.37  0.715
## Year1999         -0.1124    0.3293   -0.34  0.734
## Year2000         -0.1908    0.3211   -0.59  0.554
## Year2001         -0.2541    0.3325   -0.76  0.447
## Year2002         -0.2166    0.3967   -0.55  0.586
## Year2003         -0.2196    0.2923   -0.75  0.455
## Year2004         -0.6351    0.3985   -1.59  0.115
## Year2005         -0.2959    0.2961   -1.00  0.320
## Year2006         -0.4709    0.3126   -1.51  0.135
## Year2007         -0.4951    0.2953   -1.68  0.097 .
## Year2008         -0.4647    0.3056   -1.52  0.132
## Year2009         -0.6355    0.5596   -1.14  0.259
## Year2010         -0.2337    0.3232   -0.72  0.472
## Year2011         -0.2056    0.3041   -0.68  0.501
## Year2012         -0.5095    0.2952   -1.73  0.088 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.217, Adjusted R-squared:  0.0669
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0294 0.8480 0.9430 0.8710 0.9790 0.9980
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      9.35e-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: 107"
## [1] ""

```

```

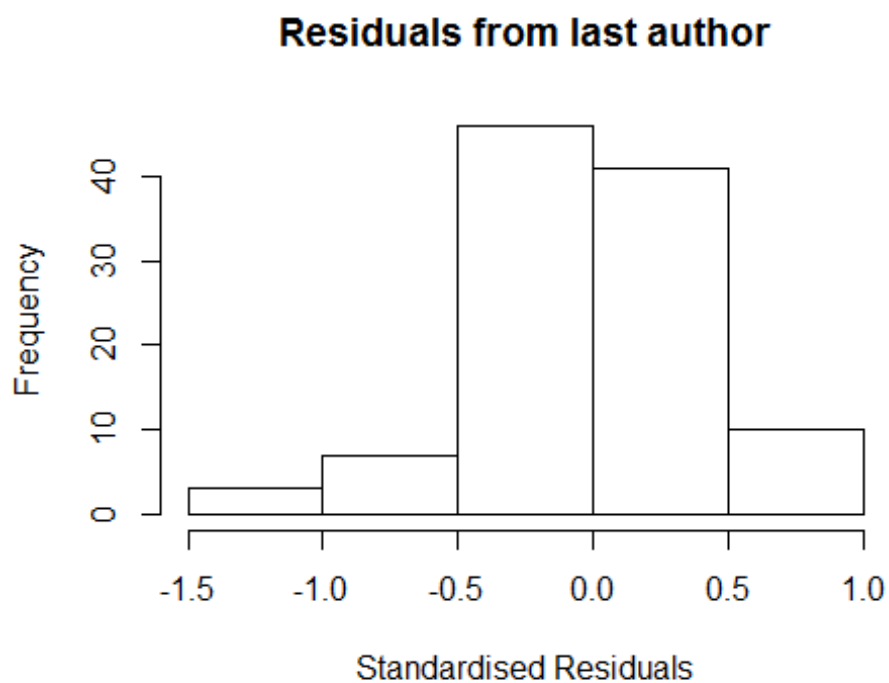
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2306"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    9    7   13   14    8   12    4    9    4    8   15    9   14   14
## 2011 2012
##   12   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    3    8    7    2    7    4    6    3    4    9    5   10    7
## 2011 2012
##   10   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    3    4    7    2    5    4    6    2    4    9    5    9    6
## 2011 2012
##    9   13
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.04767251107922"
## [1] "Male first author team size 2018 geometric mean: 3.41995189335339"

## 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 = 9, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 3.38014889649687"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

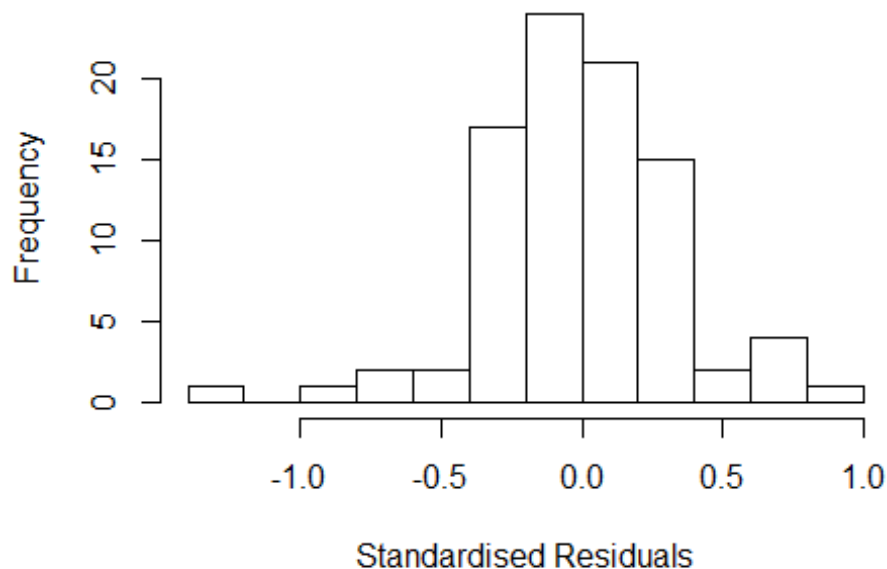
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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	1.065e+14	1	1.032e+07
## LastAuthorFemale	3.354e+00	1	1.831e+00
## UniqueAuthors	3.873e+15	4	8.882e+01
## Year	1.973e+16	16	3.230e+00

Residuals from first and last author and team size



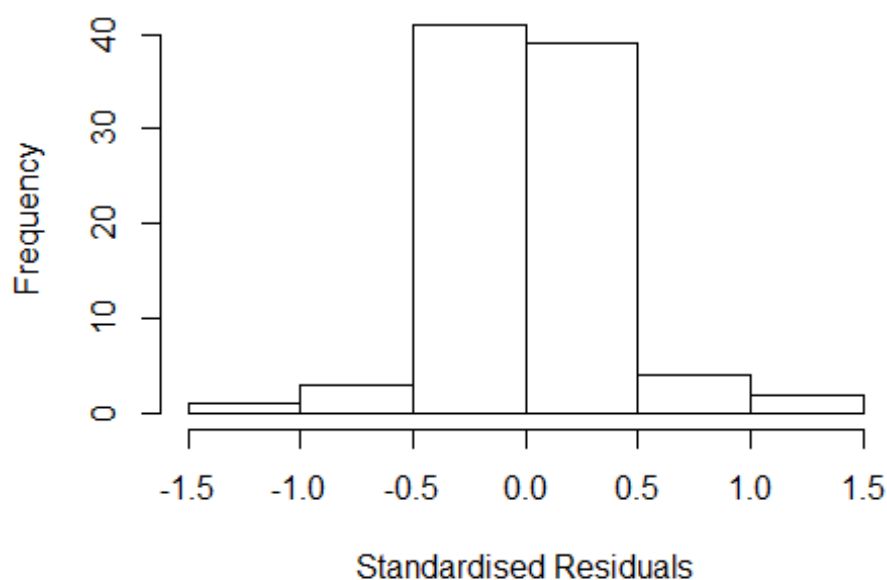
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2513 -0.1994 -0.0017  0.1804  0.9037
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7721     0.1502   11.80 < 2e-16 ***
## FirstAuthorFemale1 -0.0179     0.1150    -0.16  0.87687
## LastAuthorFemale1 -0.0158     0.1023    -0.15  0.87762
## UniqueAuthors2    -0.0473     0.1077    -0.44  0.66207
## UniqueAuthors3     0.0856     0.1237     0.69  0.49166
## UniqueAuthors4     0.3213     0.1410     2.28  0.02591 *
## UniqueAuthors5     0.2148     0.1384     1.55  0.12539
## Year1997          -1.0391     0.1502    -6.92  2.1e-09 ***
## Year1998          -0.5896     0.2344    -2.51  0.01431 *
## Year1999          -0.7323     0.1860    -3.94  0.00020 ***
```

```

## Year2000          -0.5209      0.1823    -2.86   0.00568 **
## Year2001          -0.9901      0.3213    -3.08   0.00299 **
## Year2002          -0.9174      0.1930    -4.75   1.1e-05 ***
## Year2003          -0.7383      0.2792    -2.64   0.01019 *
## Year2004          -0.5040      0.1353    -3.72   0.00040 ***
## Year2005          -0.2112      0.2748    -0.77   0.44493
## Year2006          -0.6278      0.1612    -3.89   0.00023 ***
## Year2007          -0.7165      0.1729    -4.14   9.8e-05 ***
## Year2008          -0.5861      0.1911    -3.07   0.00313 **
## Year2009          -0.8740      0.1897    -4.61   1.9e-05 ***
## Year2010          -0.7062      0.1527    -4.63   1.8e-05 ***
## Year2011          -0.8647      0.2101    -4.12   0.00011 ***
## Year2012          -1.0617      0.1652    -6.43   1.6e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.38,   Adjusted R-squared:  0.177
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0966 0.8970 0.9590 0.9020 0.9900 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          1.11e-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 6.499e+12 1      2.549e+06
## LastAuthorFemale  3.308e+00 1      1.819e+00
## Year              1.975e+13 16      2.603e+00

```


Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.2903 -0.1847 0.0037 0.1927 1.1537
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.68251 0.10221 16.46 < 2e-16 ***
## FirstAuthorFemale1 0.02449 0.10221 0.24 0.81133
## LastAuthorFemale1 -0.00366 0.10773 -0.03 0.97296
## Year1997 -0.94951 0.10221 -9.29 6.8e-14 ***
## Year1998 -0.35392 0.18147 -1.95 0.05508 .
## Year1999 -0.57657 0.16252 -3.55 0.00069 ***
## Year2000 -0.39218 0.15310 -2.56 0.01254 *
## Year2001 -0.69701 0.41181 -1.69 0.09493 .
## Year2002 -0.82326 0.17262 -4.77 9.6e-06 ***
## Year2003 -0.68933 0.28518 -2.42 0.01821 *
## Year2004 -0.35412 0.10594 -3.34 0.00133 **
## Year2005 -0.11209 0.19756 -0.57 0.57225
```

```

## Year2006      -0.55304    0.14421   -3.83  0.00027 ***
## Year2007      -0.45831    0.10453   -4.38  3.9e-05 ***
## Year2008      -0.36235    0.13030   -2.78  0.00694 **
## Year2009      -0.70197    0.14151   -4.96  4.6e-06 ***
## Year2010      -0.52551    0.14516   -3.62  0.00055 ***
## Year2011      -0.64134    0.14241   -4.50  2.6e-05 ***
## Year2012      -1.00717    0.16546   -6.09  5.3e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.336, Adjusted R-squared:  0.168
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.130  0.890  0.962  0.898  0.984  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      1.11e-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 9.007e+15  1      9.491e+07
## Year      9.007e+15 16      3.152e+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 + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max

```

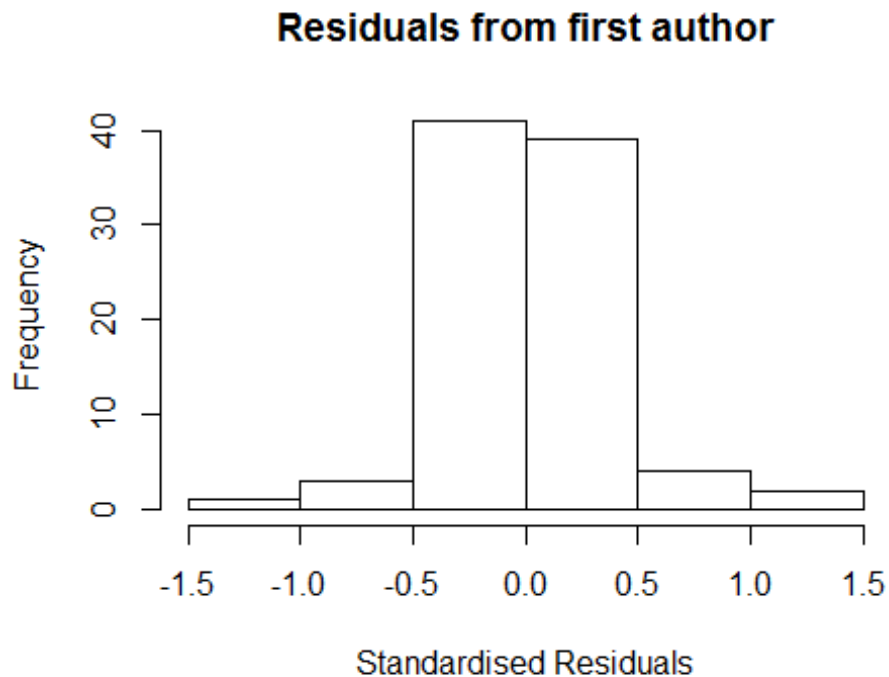
```

## -1.2916 -0.1830 0.0037 0.1927 1.1559
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.6828    0.0979   17.19 < 2e-16 ***
## FirstAuthorFemale1 0.0242    0.0979    0.25 0.80520
## Year1997         -0.9498    0.0979   -9.70 1.0e-14 ***
## Year1998         -0.3542    0.1789   -1.98 0.05156 .
## Year1999         -0.5787    0.1467   -3.94 0.00018 ***
## Year2000         -0.3912    0.1441   -2.71 0.00829 **
## Year2001         -0.6973    0.4135   -1.69 0.09611 .
## Year2002         -0.8233    0.1722   -4.78 8.9e-06 ***
## Year2003         -0.6891    0.2809   -2.45 0.01658 *
## Year2004         -0.3547    0.1077   -3.29 0.00154 **
## Year2005         -0.1159    0.1553   -0.75 0.45806
## Year2006         -0.5540    0.1429   -3.88 0.00023 ***
## Year2007         -0.4589    0.0972   -4.72 1.1e-05 ***
## Year2008         -0.3627    0.1279   -2.84 0.00594 **
## Year2009         -0.7034    0.1093   -6.43 1.2e-08 ***
## Year2010         -0.5270    0.1230   -4.29 5.6e-05 ***
## Year2011         -0.6427    0.1375   -4.67 1.3e-05 ***
## Year2012         -1.0097    0.1677   -6.02 6.7e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.343
## Multiple R-squared: 0.338, Adjusted R-squared: 0.182
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.125 0.889 0.961 0.897 0.984 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      1.11e-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

## 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 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 = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.29429 -0.17955  0.00234  0.19074  1.14989
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.70700    0.00000      Inf < 2e-16 ***
## LastAuthorFemale1 0.00171    0.10238    0.02  0.9867
```

```

## Year1997      -0.97400      0.00000      -Inf < 2e-16 ***
## Year1998      -0.37839      0.14917      -2.54  0.0134 *
## Year1999      -0.60381      0.11848      -5.10  2.7e-06 ***
## Year2000      -0.41271      0.10019      -4.12  0.0001 ***
## Year2001      -0.72150      0.40860      -1.77  0.0817 .
## Year2002      -0.83228      0.17026      -4.89  6.0e-06 ***
## Year2003      -0.71082      0.23123      -3.07  0.0030 **
## Year2004      -0.37468      0.07254      -5.17  2.1e-06 ***
## Year2005      -0.12971      0.18837      -0.69  0.4933
## Year2006      -0.57268      0.12316      -4.65  1.5e-05 ***
## Year2007      -0.47386      0.09495      -4.99  4.0e-06 ***
## Year2008      -0.38703      0.08571      -4.52  2.4e-05 ***
## Year2009      -0.71436      0.13690      -5.22  1.7e-06 ***
## Year2010      -0.54747      0.10307      -5.31  1.2e-06 ***
## Year2011      -0.65824      0.13085      -5.03  3.5e-06 ***
## Year2012      -1.02789      0.17539      -5.86  1.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.339
## Multiple R-squared:  0.34,   Adjusted R-squared:  0.184
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 78 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.113  0.886  0.964  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           1.11e-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: 90"
## [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
##   15   18   34   23   13   12   11   19    8   11    8   14   10   12   17
## 2011 2012
##   10   18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    3    8   13    4    1    7   11    5    8    4    9    6   10   11
## 2011 2012
##    6   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    2    8   12    4    0    6    9    5    6    3    9    6    8   11
## 2011 2012
##    6   14
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.42961136299083"
## [1] "Male first author team size 2018 geometric mean: 5.09145979004366"

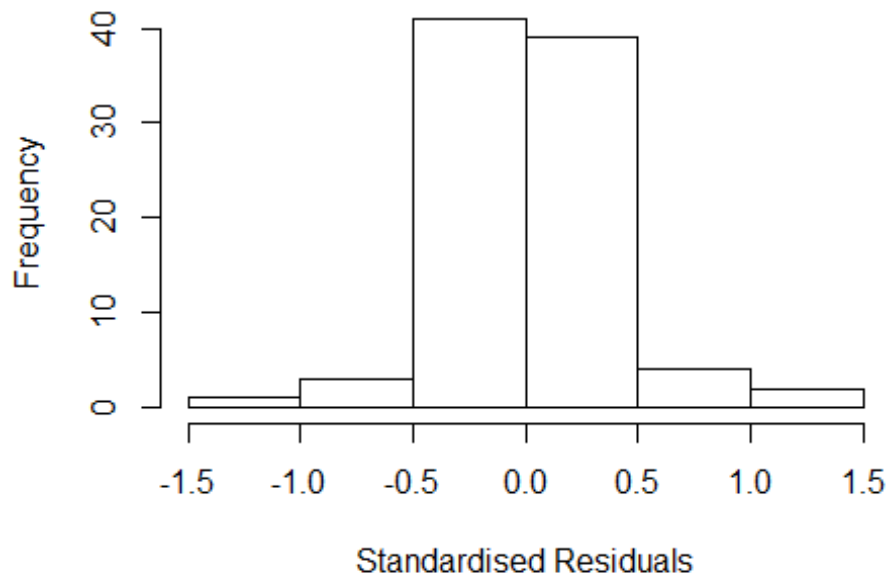
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.85961686300321"
## [1] "Male last author team size 2018 geometric mean: 4.34877739291784"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

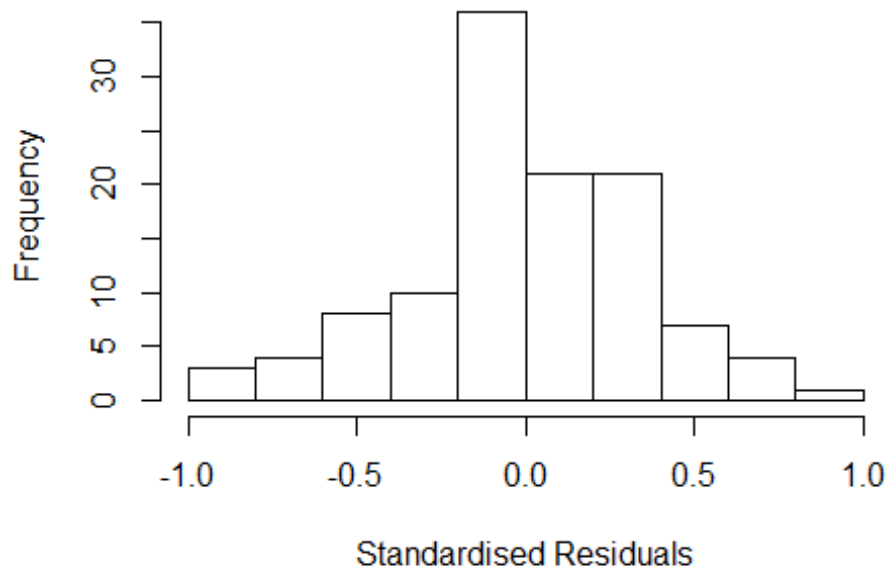
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 48, 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  1.678  1      1.295
## LastAuthorFemale   2.730  1      1.652
## UniqueAuthors     52.473  4      1.641
## Year              126.647 15      1.175
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -0.89452 -0.18748 -0.00473 0.23878 0.92536
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.046570 0.190707 5.49 3.5e-07 ***
## FirstAuthorFemale1 -0.041598 0.068336 -0.61 0.5442
## LastAuthorFemale1 0.123898 0.093480 1.33 0.1883
## UniqueAuthors2 0.272803 0.159561 1.71 0.0907 .
## UniqueAuthors3 0.167664 0.200614 0.84 0.4054
## UniqueAuthors4 0.106761 0.183247 0.58 0.5616
## UniqueAuthors5 0.198602 0.146892 1.35 0.1796
## Year1997 0.387629 0.163542 2.37 0.0198 *
## Year1998 0.030221 0.165388 0.18 0.8554
## Year1999 0.003365 0.141450 0.02 0.9811
```

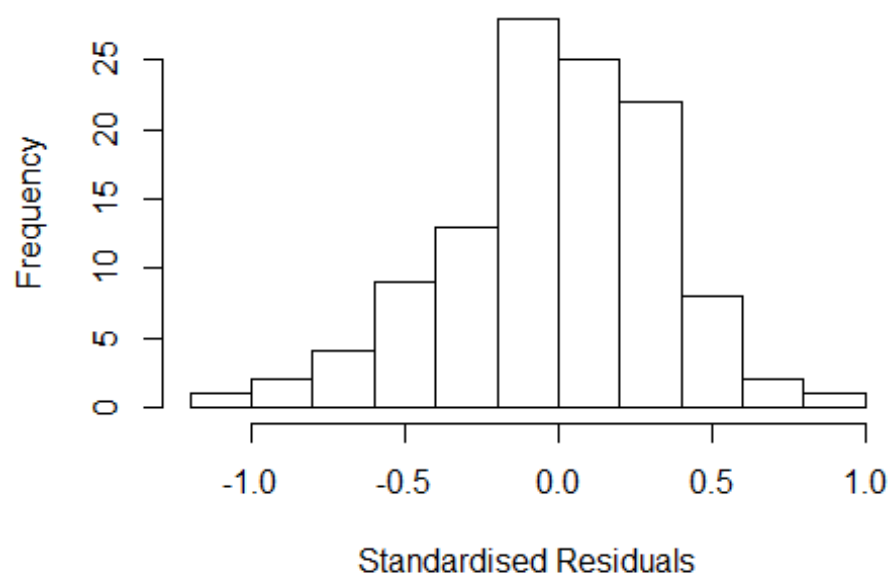


```

## Year2000      -0.571352    0.193522    -2.95    0.0040 **
## Year2002      -0.058273    0.228527    -0.25    0.7993
## Year2003      -0.258814    0.271363    -0.95    0.3427
## Year2004       0.016893    0.407322     0.04    0.9670
## Year2005      -0.051096    0.146933    -0.35    0.7288
## Year2006      -0.338651    0.237400    -1.43    0.1571
## Year2007      -0.100096    0.137328    -0.73    0.4679
## Year2008      -0.090885    0.195281    -0.47    0.6427
## Year2009      -0.371591    0.141711    -2.62    0.0102 *
## Year2010       0.000461    0.186680     0.00    0.9980
## Year2011      -0.484244    0.143542    -3.37    0.0011 **
## Year2012      -0.033844    0.131221    -0.26    0.7970
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.291, Adjusted R-squared:  0.131
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 104 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.449  0.876  0.959  0.908  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      8.70e-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 1.346 1 1.160
## LastAuthorFemale 1.623 1 1.274
## Year 2.167 15 1.026

```

Residuals from first and last author



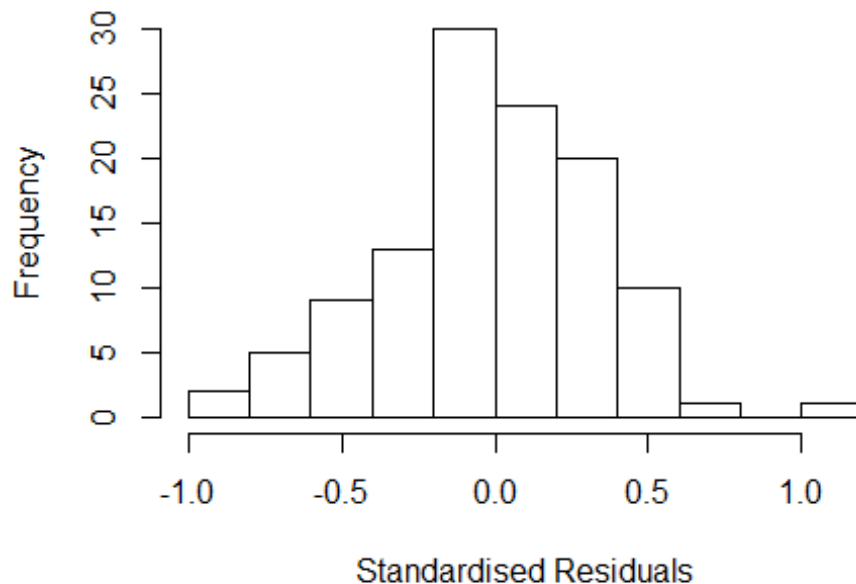
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.05887 -0.20025 0.00559 0.24117 0.97699
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.26827 0.11372 11.15 < 2e-16 ***
## FirstAuthorFemale1 -0.05656 0.06823 -0.83 0.40912
## LastAuthorFemale1 0.13451 0.07918 1.70 0.09257 .
## Year1997 0.26523 0.17285 1.53 0.12817
## Year1998 -0.02329 0.18750 -0.12 0.90139
## Year1999 -0.00777 0.16514 -0.05 0.96256
## Year2000 -0.58209 0.17364 -3.35 0.00114 **
## Year2002 -0.10451 0.20020 -0.52 0.60285
## Year2003 -0.28735 0.28445 -1.01 0.31493
## Year2004 -0.00837 0.34441 -0.02 0.98067
## Year2005 -0.11773 0.16633 -0.71 0.48076
## Year2006 -0.39174 0.14019 -2.79 0.00627 **
```

```

## Year2007      -0.15627    0.12997   -1.20  0.23217
## Year2008      -0.10419    0.22817   -0.46  0.64897
## Year2009      -0.48739    0.12463   -3.91  0.00017 ***
## Year2010      -0.04969    0.15989   -0.31  0.75665
## Year2011      -0.49623    0.13177   -3.77  0.00028 ***
## Year2012      -0.03405    0.12523   -0.27  0.78627
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.349
## Multiple R-squared:  0.257, Adjusted R-squared:  0.127
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 105 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.337  0.883  0.955  0.907  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      8.70e-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 1.401 1      1.184
## Year              1.401 15      1.011

```

Residuals from first author



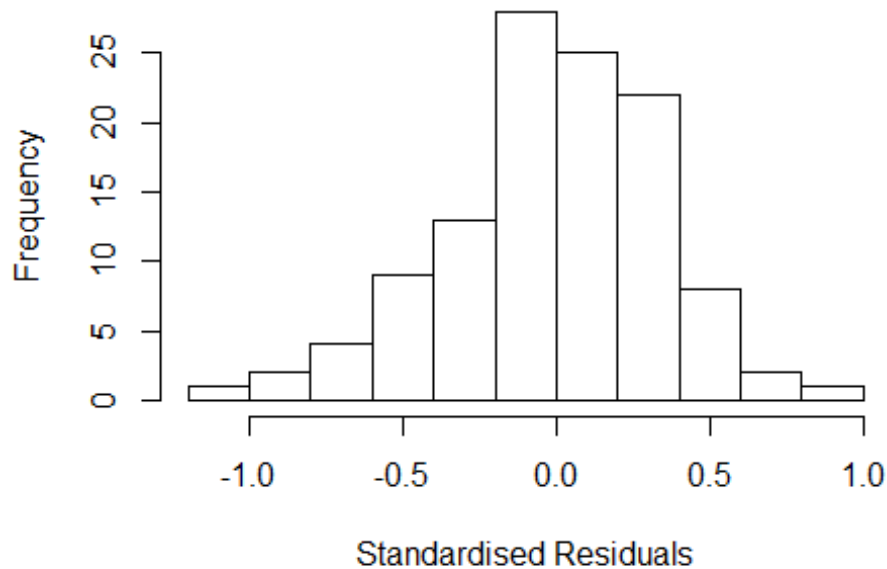
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.98559 -0.19979 -0.00145  0.22789  1.03808
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.280091   0.111723   11.46  < 2e-16 ***
## FirstAuthorFemale1 -0.031472   0.073620   -0.43  0.66996
## Year1997         0.253409   0.171646    1.48  0.14306
## Year1998         0.000768   0.198362    0.00  0.99692
## Year1999         0.053827   0.170698    0.32  0.75318
## Year2000        -0.572767   0.195420   -2.93  0.00421 **
## Year2002        -0.036152   0.181314   -0.20  0.84237
## Year2003        -0.294503   0.286597   -1.03  0.30667
## Year2004        -0.019925   0.346454   -0.06  0.95426
## Year2005        -0.123196   0.154825   -0.80  0.42812
## Year2006        -0.364986   0.159845   -2.28  0.02457 *
## Year2007        -0.126653   0.129577   -0.98  0.33076
```

```

## Year2008          -0.096850    0.230083    -0.42    0.67472
## Year2009          -0.455415    0.120747    -3.77    0.00028 ***
## Year2010          -0.007324    0.165170    -0.04    0.96472
## Year2011          -0.494497    0.125737    -3.93    0.00016 ***
## Year2012          -0.032832    0.124116    -0.26    0.79193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.237, Adjusted R-squared:  0.113
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 100 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.346  0.854  0.946  0.898  0.985  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          8.70e-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 1.702 1          1.305
## Year            1.702 15          1.018

```

Residuals from last author



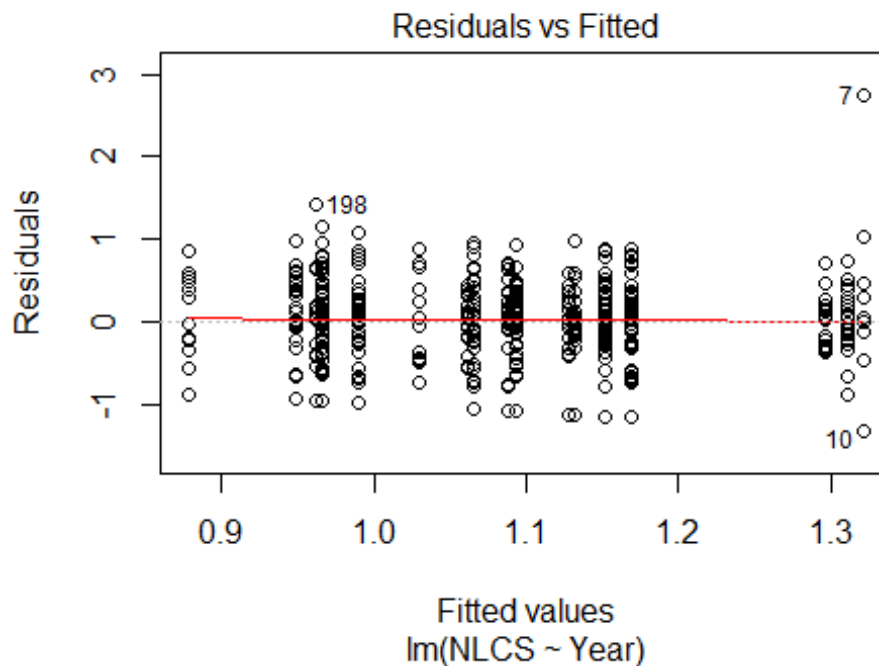
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.07309 -0.20352 0.00663 0.21826 1.00631
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.24070 0.10434 11.89 < 2e-16 ***
## LastAuthorFemale1 0.12120 0.08471 1.43 0.15566
## Year1997 0.29280 0.16664 1.76 0.08203 .
## Year1998 -0.02376 0.18508 -0.13 0.89812
## Year1999 0.00379 0.16135 0.02 0.98129
## Year2000 -0.56380 0.16187 -3.48 0.00074 ***
## Year2002 -0.07814 0.19908 -0.39 0.69552
## Year2003 -0.28880 0.27982 -1.03 0.30457
## Year2004 0.01875 0.33568 0.06 0.95556
## Year2005 -0.11313 0.16857 -0.67 0.50372
## Year2006 -0.37724 0.14968 -2.52 0.01334 *
## Year2007 -0.13609 0.12112 -1.12 0.26393
```

```

## Year2008          -0.09070      0.23293    -0.39  0.69783
## Year2009          -0.47653      0.12513    -3.81  0.00024 ***
## Year2010          -0.03910      0.15620    -0.25  0.80285
## Year2011          -0.48556      0.13341    -3.64  0.00044 ***
## Year2012          -0.03780      0.12724    -0.30  0.76706
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.251, Adjusted R-squared:  0.129
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 104 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.342  0.890  0.953  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      8.70e-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: 115"
## [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
##   23   27   19   32   36   34   36   46   35   37   48   35   48   61   61
## 2011 2012
##   68   83
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11   18   13   16   13   17   21   34   22   26   35   24   32   43   46
## 2011 2012
##   53   66

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   16   12   13   12   17   20   32   19   25   31   22   29   36   39
## 2011 2012
##   50   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 = 56, df = 16, p-value = 2e-06
```

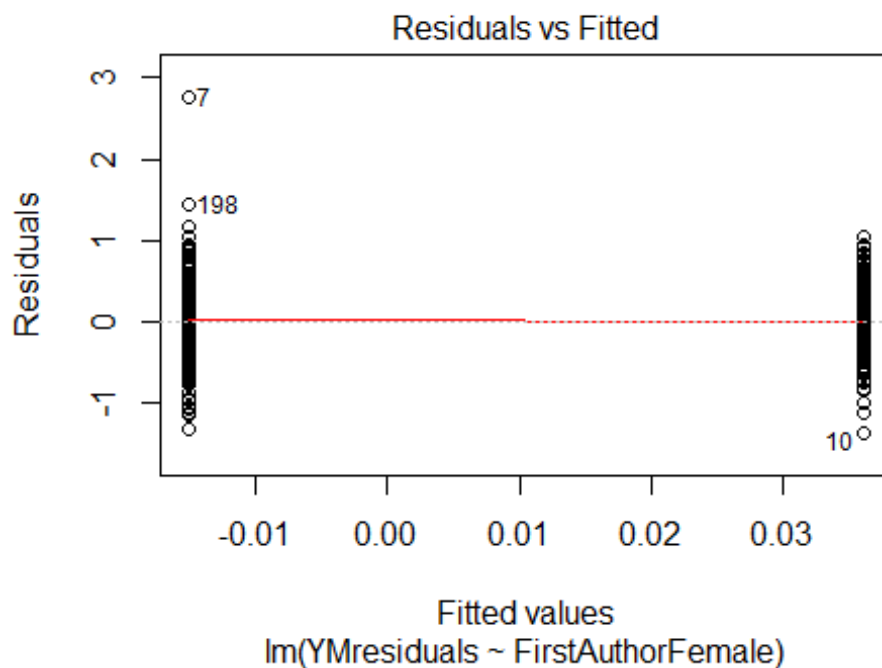


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.61, df = 1, p-value = 0.4
## [1] "Female first author team size 2018 geometric mean: 2.53415106527303"
## [1] "Male first author team size 2018 geometric mean: 2.82821996667806"
## 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] "Female last author team size 2018 geometric mean: 2.60397946182773"
## [1] "Male last author team size 2018 geometric mean: 2.78754532928041"

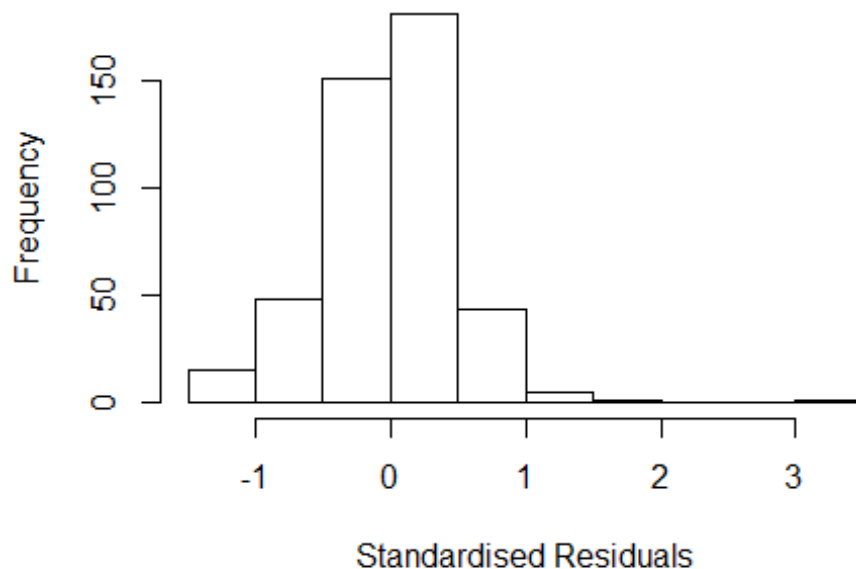
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 310, 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	1.601	1	1.265
LastAuthorFemale	1.448	1	1.203
UniqueAuthors	2.140	4	1.100
Year	3.048	16	1.035

Residuals from first and last author and team size



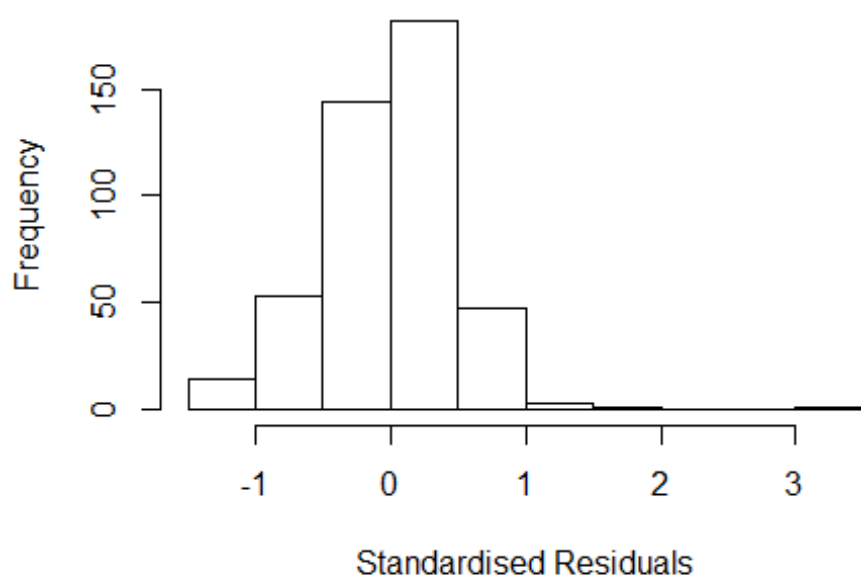
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 7 0030135572 4.064 1996      2100      2      3.242
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2901 -0.2810  0.0221  0.2898  3.2423
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8217    0.5788    1.42   0.1565
## FirstAuthorFemale1 0.0718    0.0565    1.27   0.2038
## LastAuthorFemale1 -0.0585    0.0578   -1.01   0.3125
## UniqueAuthors2    0.1945    0.0591    3.29   0.0011 **
## UniqueAuthors3    0.1767    0.0741    2.39   0.0175 *
## UniqueAuthors4    0.1625    0.0729    2.23   0.0264 *
## UniqueAuthors5    0.2775    0.0702    3.95 9.1e-05 ***
## Year1997          0.2586    0.5808    0.45   0.6564
## Year1998          0.0348    0.5885    0.06   0.9529
## Year1999          0.4109    0.5809    0.71   0.4798
```

```

## Year2000          0.0315      0.5956      0.05      0.9578
## Year2001          0.0325      0.6009      0.05      0.9570
## Year2002          0.3662      0.5761      0.64      0.5253
## Year2003          0.0937      0.5811      0.16      0.8720
## Year2004          0.1208      0.5765      0.21      0.8342
## Year2005          0.2036      0.5817      0.35      0.7265
## Year2006          0.0325      0.5804      0.06      0.9553
## Year2007          0.1747      0.5803      0.30      0.7635
## Year2008          0.0818      0.5872      0.14      0.8892
## Year2009          0.1833      0.5785      0.32      0.7515
## Year2010          0.2740      0.5770      0.47      0.6352
## Year2011          0.2049      0.5755      0.36      0.7220
## Year2012         -0.0141      0.5784     -0.02      0.9806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.118, Adjusted R-squared:  0.0716
## Convergence in 40 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.00022);
## 40 weights are ~= 1. The remaining 404 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.143  0.849  0.947  0.889  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-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 1.496 1          1.223
## LastAuthorFemale  1.479 1          1.216
## Year              1.623 16          1.015

```

Residuals from first and last author

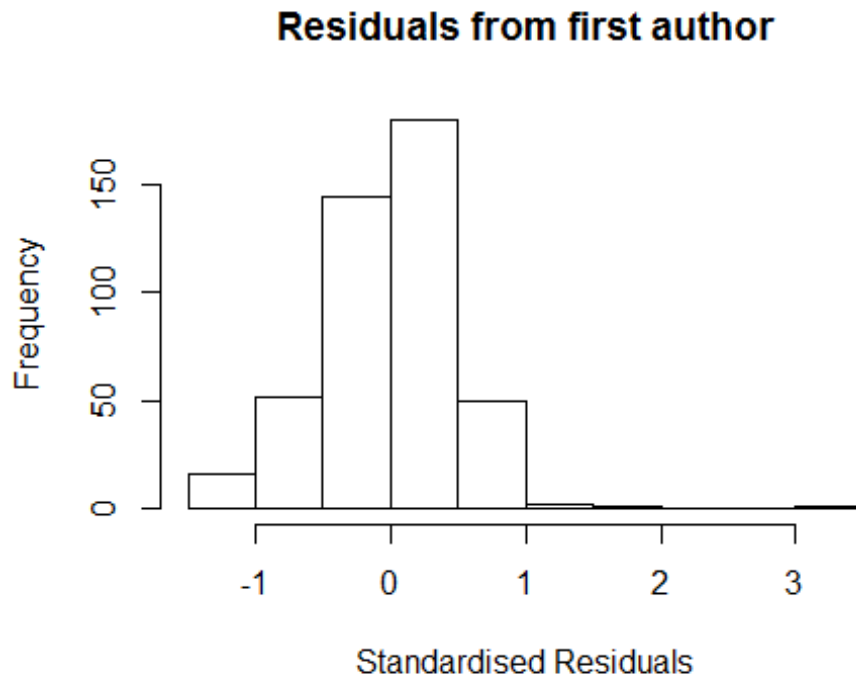


```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 7 0030135572 4.064 1996      2100      2      3.067
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2121 -0.2697  0.0166  0.2646  3.0675
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.99653    0.57698   1.73   0.085 .
## FirstAuthorFemale1 0.09324    0.05596   1.67   0.096 .
## LastAuthorFemale1 -0.07117    0.05982  -1.19   0.235
## Year1997         0.18527    0.58570   0.32   0.752
## Year1998        -0.07836    0.59818  -0.13   0.896
## Year1999         0.34911    0.58530   0.60   0.551
## Year2000        -0.03997    0.61439  -0.07   0.948
## Year2001        -0.10789    0.60077  -0.18   0.858
## Year2002         0.34284    0.58432   0.59   0.558
## Year2003         0.06535    0.59129   0.11   0.912
## Year2004         0.07131    0.58251   0.12   0.903
## Year2005         0.15471    0.58641   0.26   0.792
```

```

## Year2006      -0.03810    0.58697   -0.06    0.948
## Year2007      0.12154    0.58372    0.21    0.835
## Year2008     -0.00108    0.59231    0.00    0.999
## Year2009      0.11794    0.58554    0.20    0.840
## Year2010      0.21561    0.58806    0.37    0.714
## Year2011      0.16916    0.58378    0.29    0.772
## Year2012     -0.05029    0.58532   -0.09    0.932
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.411
## Multiple R-squared:  0.0752, Adjusted R-squared:  0.0361
## Convergence in 42 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.00022);
## 40 weights are ~= 1. The remaining 404 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.152  0.833   0.951   0.883   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-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 1.331 1          1.154
## Year              1.331 16          1.009

```



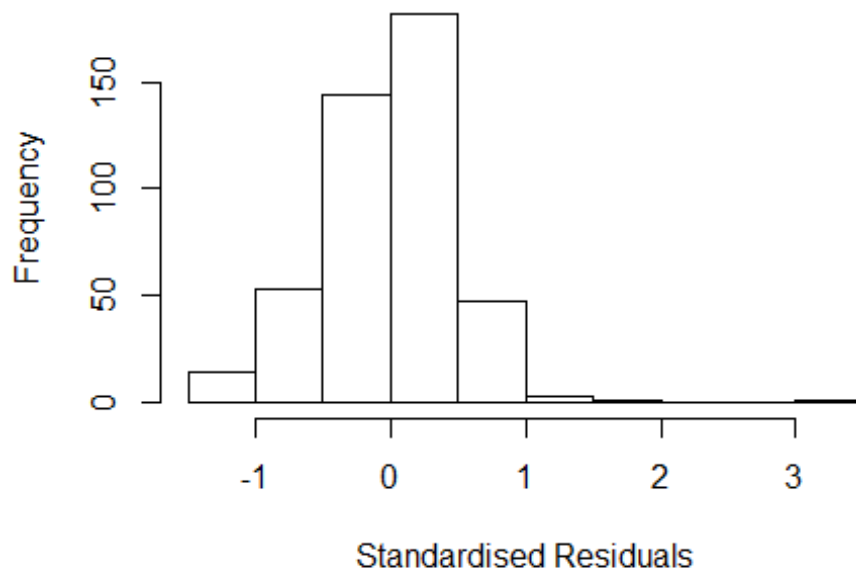
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 7 0030135572 4.064 1996      2100      2      3.067
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2056 -0.2722  0.0251  0.2695  3.0597
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0043     0.6040   1.66   0.097 .
## FirstAuthorFemale1  0.0626     0.0530   1.18   0.238
## Year1997          0.1824     0.6112   0.30   0.766
## Year1998         -0.0828     0.6260  -0.13   0.895
## Year1999          0.3387     0.6123   0.55   0.580
## Year2000         -0.0471     0.6377  -0.07   0.941
## Year2001         -0.1209     0.6273  -0.19   0.847
## Year2002          0.3302     0.6103   0.54   0.589
## Year2003          0.0443     0.6144   0.07   0.943
## Year2004          0.0673     0.6092   0.11   0.912
## Year2005          0.1387     0.6100   0.23   0.820
## Year2006         -0.0544     0.6126  -0.09   0.929
```

```

## Year2007          0.1003      0.6094      0.16      0.869
## Year2008         -0.0220      0.6175     -0.04      0.972
## Year2009          0.1013      0.6103      0.17      0.868
## Year2010          0.1977      0.6127      0.32      0.747
## Year2011          0.1417      0.6086      0.23      0.816
## Year2012         -0.0683      0.6111     -0.11      0.911
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.072, Adjusted R-squared:  0.0351
## Convergence in 41 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.00022);
## 38 weights are ~= 1. The remaining 406 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.145  0.831  0.951  0.882  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      2.25e-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 1.299 1          1.140
## Year              1.299 16          1.008

```

Residuals from last author



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 7 0030135572 4.064 1996      2100      2      3.067
##
## Call:
## lmrob(formula = NLCS ~ 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 -0.2573  0.0197  0.2715  3.0721
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9919     0.5664   1.75   0.081 .
## LastAuthorFemale1 -0.0265     0.0560  -0.47   0.636
## Year1997          0.2040     0.5727   0.36   0.722
## Year1998         -0.0628     0.5884  -0.11   0.915
## Year1999          0.3687     0.5736   0.64   0.521
## Year2000         -0.0347     0.6028  -0.06   0.954
## Year2001         -0.0790     0.5892  -0.13   0.893
## Year2002          0.3536     0.5729   0.62   0.537
## Year2003          0.0965     0.5769   0.17   0.867
## Year2004          0.0873     0.5709   0.15   0.878
## Year2005          0.1706     0.5734   0.30   0.766
## Year2006         -0.0116     0.5751  -0.02   0.984
```

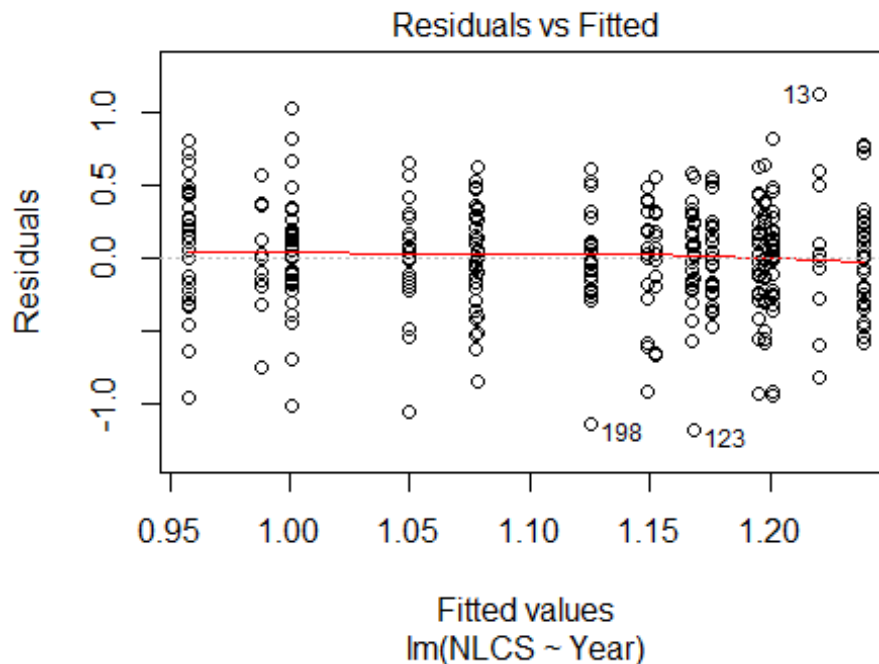


```

## Year2007          0.1346      0.5726      0.24      0.814
## Year2008          0.0244      0.5795      0.04      0.966
## Year2009          0.1381      0.5727      0.24      0.810
## Year2010          0.2480      0.5741      0.43      0.666
## Year2011          0.1824      0.5721      0.32      0.750
## Year2012         -0.0204      0.5719     -0.04      0.972
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.409
## Multiple R-squared:  0.0691, Adjusted R-squared:  0.032
## Convergence in 38 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.00022);
## 38 weights are ~= 1. The remaining 406 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.164  0.829  0.952  0.882  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      2.25e-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: 445"
## [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
##   22   25   16   32   21   27   35   34   31   18   27   27   29   29   47
## 2011 2012
##   44   47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11   12   11   22    9   17   23   24   21   12   21   23   25   25   31

```

```
## 2011 2012
## 30 44
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 12 10 18 9 16 22 23 17 12 19 21 20 21 27
## 2011 2012
## 29 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 = 19, df = 16, p-value = 0.2
```



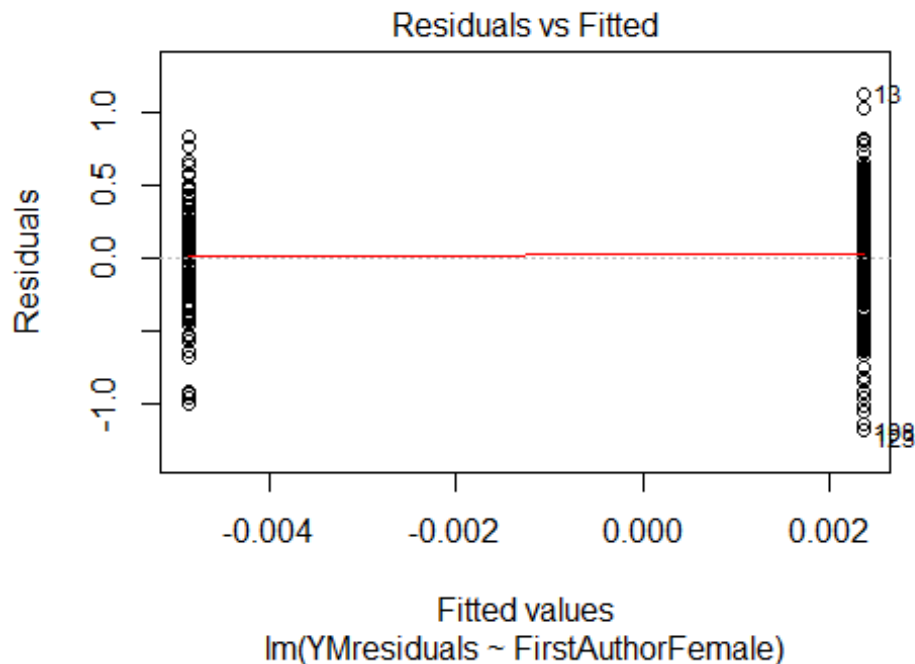
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.5, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 2.16093273781394"
## [1] "Male first author team size 2018 geometric mean: 2.86621584553468"

## 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 = 88, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.18857476102454"
## [1] "Male last author team size 2018 geometric mean: 2.7536287387554"

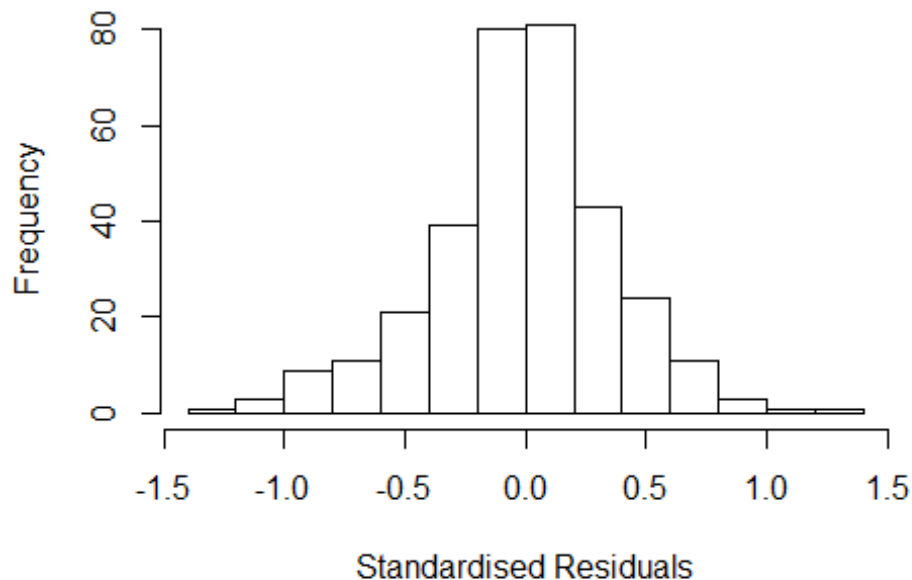
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 83, 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	1.418	1	1.191
LastAuthorFemale	1.364	1	1.168
UniqueAuthors	3.203	4	1.157
Year	5.266	16	1.053

Residuals from first and last author and team size



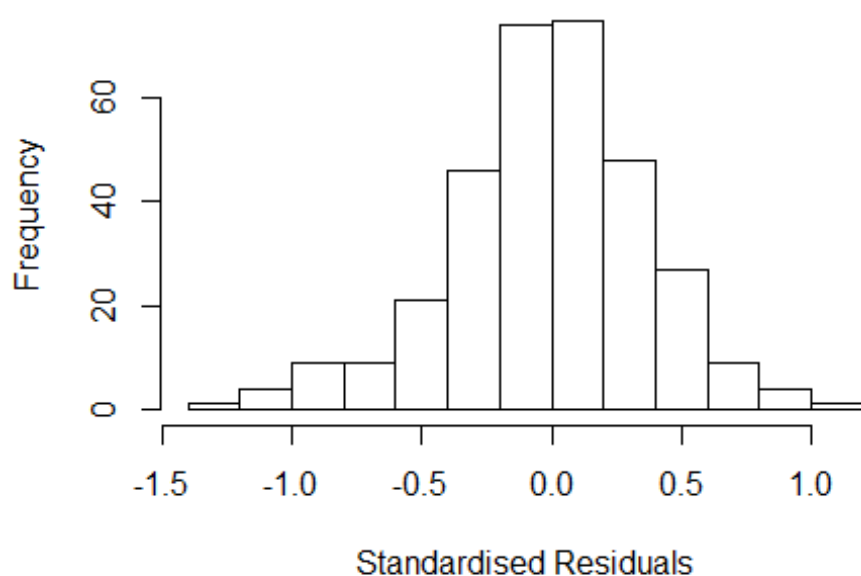
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28696 -0.21187 -0.00417  0.20352  1.22787
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.11713    0.23324   4.79 2.6e-06 ***
## FirstAuthorFemale1  0.01226    0.04466   0.27  0.784
## LastAuthorFemale1 -0.08404    0.05455  -1.54  0.124
## UniqueAuthors2     0.00866    0.07635   0.11  0.910
## UniqueAuthors3     0.02464    0.08039   0.31  0.759
## UniqueAuthors4     0.11790    0.08714   1.35  0.177
## UniqueAuthors5     0.18695    0.08779   2.13  0.034 *
## Year1997         -0.00920    0.23711  -0.04  0.969
## Year1998         -0.14501    0.22826  -0.64  0.526
## Year1999          0.04765    0.21500   0.22  0.825
```

```

## Year2000      0.16983    0.21762    0.78    0.436
## Year2001      0.04255    0.22607    0.19    0.851
## Year2002      0.00267    0.20907    0.01    0.990
## Year2003      0.01974    0.21030    0.09    0.925
## Year2004      0.07193    0.21522    0.33    0.738
## Year2005      0.00942    0.22361    0.04    0.966
## Year2006     -0.12775    0.21398   -0.60    0.551
## Year2007      0.01759    0.21760    0.08    0.936
## Year2008      0.02217    0.21740    0.10    0.919
## Year2009     -0.06596    0.21694   -0.30    0.761
## Year2010      0.08009    0.21367    0.37    0.708
## Year2011     -0.12831    0.21149   -0.61    0.544
## Year2012     -0.16314    0.22208   -0.73    0.463
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.339
## Multiple R-squared:  0.0847, Adjusted R-squared:  0.0187
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 40 weights are ~= 1. The remaining 288 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.119  0.850  0.953  0.884  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      3.05e-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 1.290 1      1.136
## LastAuthorFemale  1.397 1      1.182
## Year              1.712 16      1.017

```

Residuals from first and last author



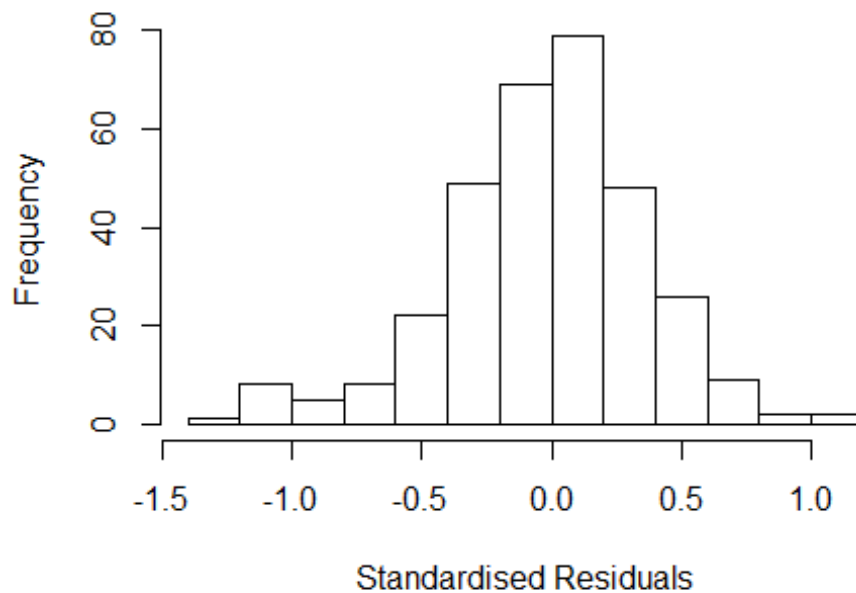
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.30620 -0.22305  0.00167  0.21899  1.18771
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.15729    0.19553     5.92 8.6e-09 ***
## FirstAuthorFemale1  0.01611    0.04456     0.36  0.72
## LastAuthorFemale1 -0.08255    0.05630    -1.47  0.14
## Year1997         -0.03407    0.22841    -0.15  0.88
## Year1998         -0.17855    0.21704    -0.82  0.41
## Year1999          0.03202    0.20405     0.16  0.88
## Year2000          0.14891    0.20790     0.72  0.47
## Year2001          0.01361    0.21825     0.06  0.95
## Year2002         -0.00288    0.20463    -0.01  0.99
## Year2003          0.05302    0.20626     0.26  0.80
## Year2004          0.06662    0.21069     0.32  0.75
## Year2005          0.02233    0.21932     0.10  0.92
```

```

## Year2006      -0.10692    0.20798   -0.51    0.61
## Year2007      0.01829    0.20564    0.09    0.93
## Year2008      0.03049    0.20992    0.15    0.88
## Year2009     -0.06280    0.20489   -0.31    0.76
## Year2010      0.08259    0.20909    0.40    0.69
## Year2011     -0.10196    0.20630   -0.49    0.62
## Year2012     -0.11950    0.21164   -0.56    0.57
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.348
## Multiple R-squared:  0.0554, Adjusted R-squared:  0.00039
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 301 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.128  0.861  0.959  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      3.05e-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 1.251 1      1.118
## Year              1.251 16      1.007

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2836 -0.2193 0.0075 0.2278 1.1976
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.147404 0.198302 5.79 1.8e-08 ***
## FirstAuthorFemale1 -0.001133 0.044752 -0.03 0.98
## Year1997 -0.036638 0.230340 -0.16 0.87
## Year1998 -0.201508 0.220556 -0.91 0.36
## Year1999 0.018129 0.207970 0.09 0.93
## Year2000 0.136209 0.214420 0.64 0.53
## Year2001 0.022102 0.221751 0.10 0.92
## Year2002 0.000549 0.207720 0.00 1.00
## Year2003 0.062042 0.208827 0.30 0.77
## Year2004 0.063546 0.212663 0.30 0.77
## Year2005 0.028747 0.222768 0.13 0.90
## Year2006 -0.102609 0.210904 -0.49 0.63
```

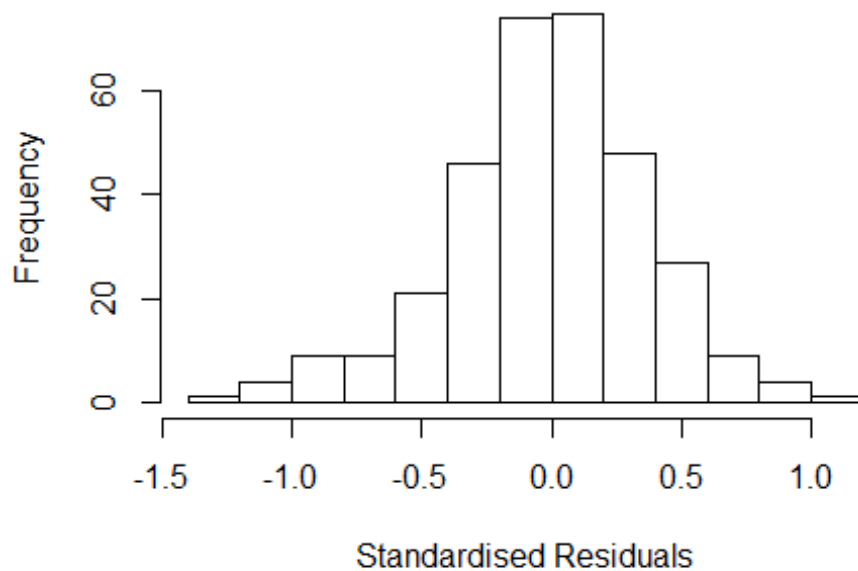


```

## Year2007          0.017183    0.207960    0.08    0.93
## Year2008          0.030306    0.213163    0.14    0.89
## Year2009         -0.060900    0.207592   -0.29    0.77
## Year2010          0.070800    0.212202    0.33    0.74
## Year2011         -0.117599    0.210142   -0.56    0.58
## Year2012         -0.121764    0.214998   -0.57    0.57
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.0464, Adjusted R-squared:  -0.00591
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 296 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.152  0.872  0.952  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      3.05e-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 1.368 1          1.17
## Year            1.368 16          1.01

```

Residuals from last author



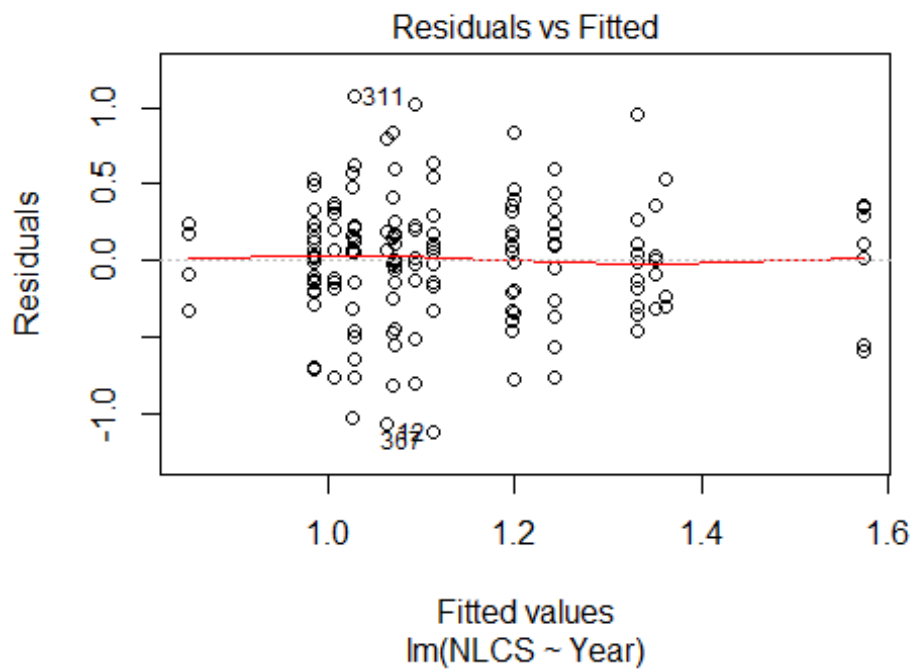
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.307849 -0.228298 -0.000464  0.217757  1.186278
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.15872    0.19610     5.91 9.1e-09 ***
## LastAuthorFemale1 -0.07809    0.05573    -1.40   0.16
## Year1997        -0.03097    0.22986    -0.13   0.89
## Year1998        -0.17626    0.21814    -0.81   0.42
## Year1999         0.03386    0.20517     0.17   0.87
## Year2000         0.14913    0.20824     0.72   0.47
## Year2001         0.01432    0.21906     0.07   0.95
## Year2002        -0.00187    0.20562    -0.01   0.99
## Year2003         0.05564    0.20769     0.27   0.79
## Year2004         0.06814    0.21192     0.32   0.75
## Year2005         0.02639    0.22158     0.12   0.91
## Year2006        -0.10291    0.20921    -0.49   0.62
```

```

## Year2007      0.02114      0.20718      0.10      0.92
## Year2008      0.03788      0.21070      0.18      0.86
## Year2009     -0.06019      0.20622     -0.29      0.77
## Year2010      0.08816      0.20966      0.42      0.67
## Year2011     -0.09867      0.20768     -0.48      0.64
## Year2012     -0.11498      0.21238     -0.54      0.59
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.347
## Multiple R-squared:  0.0549, Adjusted R-squared:  0.00312
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 300 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.123  0.863  0.958  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      3.05e-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: 328"
## [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
##   16   19   20   22   15   15   15   28   20   18   22   22   14   17   15
## 2011 2012
##   27   33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    5    7    8    4    3    7   11   12   11    8   13   10   12    9
## 2011 2012

```

```
## 13 22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 5 6 7 4 2 7 9 11 8 6 12 10 11 8
## 2011 2012
## 10 22
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 19, df = 16, p-value = 0.3
```



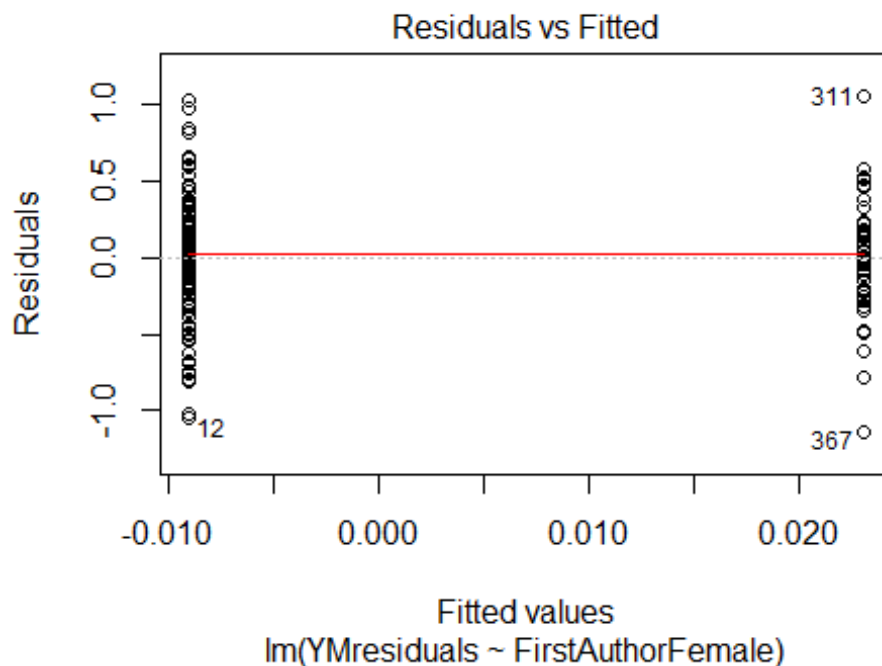
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.16, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 3.50482680803767"
## [1] "Male first author team size 2018 geometric mean: 4.03418529418011"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

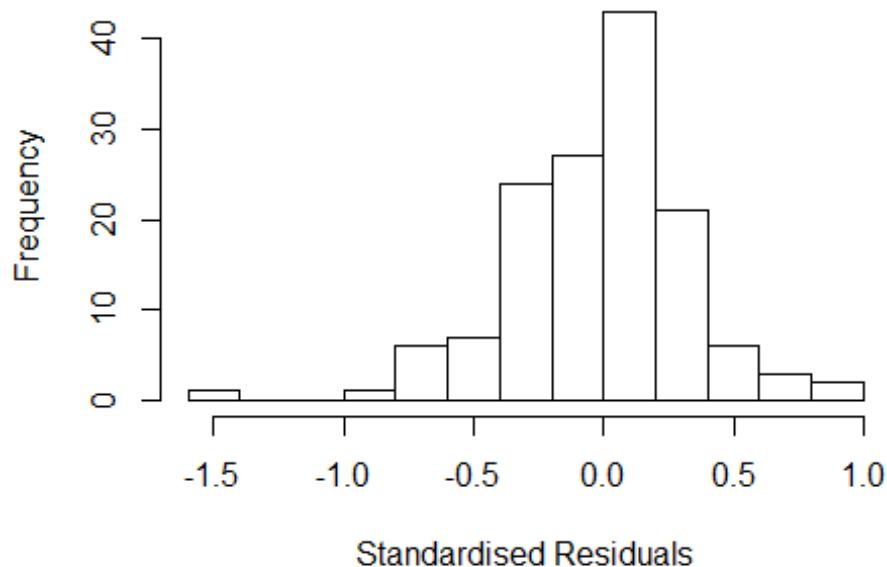
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 92, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.24534222319921"
## [1] "Male last author team size 2018 geometric mean: 3.85086036573409"

## 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 = 52, 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  3.992  1      1.998
## LastAuthorFemale   6.399  1      2.530
## UniqueAuthors    74.254  4      1.713
## Year             224.013 16      1.184
```

Residuals from first and last author and team size



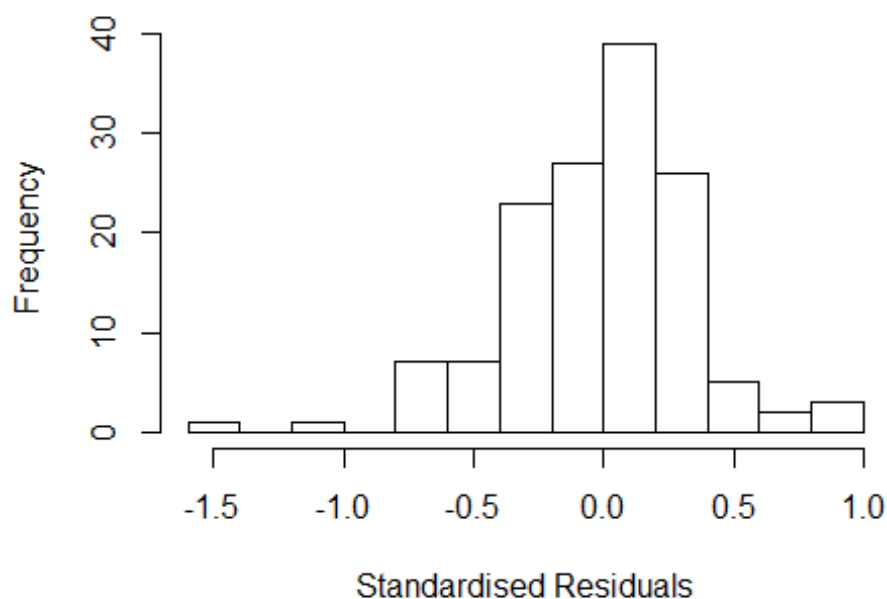
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5491 -0.2143 0.0318 0.1816 0.8604
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3497 0.7470 1.81 0.073 .
## FirstAuthorFemale1 0.0922 0.0726 1.27 0.207
## LastAuthorFemale1 0.1112 0.0904 1.23 0.221
## UniqueAuthors2 0.1994 0.1468 1.36 0.177
## UniqueAuthors3 0.1407 0.1375 1.02 0.308
## UniqueAuthors4 0.1439 0.1443 1.00 0.321
## UniqueAuthors5 0.2335 0.1304 1.79 0.076 .
## Year1997 -0.1468 0.7501 -0.20 0.845
## Year1998 0.0918 0.7754 0.12 0.906
## Year1999 -0.4107 0.7625 -0.54 0.591
```

```

## Year2000      -0.7283      0.7506      -0.97      0.334
## Year2001      -0.4576      0.7489      -0.61      0.542
## Year2002      -0.4394      0.7647      -0.57      0.567
## Year2003      -0.4112      0.7452      -0.55      0.582
## Year2004      -0.3669      0.7546      -0.49      0.628
## Year2005      -0.3567      0.7485      -0.48      0.635
## Year2006      -0.4774      0.7661      -0.62      0.534
## Year2007      -0.5424      0.7493      -0.72      0.471
## Year2008      -0.5478      0.7485      -0.73      0.466
## Year2009      -0.4363      0.7575      -0.58      0.566
## Year2010      -0.3145      0.7685      -0.41      0.683
## Year2011      -0.4334      0.7469      -0.58      0.563
## Year2012      -0.5575      0.7444      -0.75      0.455
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.325
## Multiple R-squared:  0.201, Adjusted R-squared:  0.0516
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| = 0 ( < 0.00071);
## 14 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.413  0.895  0.958  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           7.09e-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 3.979 1 1.995
## LastAuthorFemale 5.733 1 2.394
## Year 13.774 16 1.085

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.4349 -0.2368  0.0107  0.2103  0.8941
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4349     0.6720   2.14   0.035 *
## FirstAuthorFemale1  0.0878     0.0770   1.14   0.256
## LastAuthorFemale1  0.1162     0.0900   1.29   0.199
## Year1997          -0.1059     0.6774  -0.16   0.876
## Year1998           0.1044     0.6945   0.15   0.881
## Year1999          -0.3280     0.6881  -0.48   0.634
## Year2000          -0.6450     0.6662  -0.97   0.335
## Year2001          -0.5434     0.6544  -0.83   0.408
## Year2002          -0.3385     0.6730  -0.50   0.616
## Year2003          -0.3566     0.6670  -0.53   0.594
## Year2004          -0.2546     0.6713  -0.38   0.705
## Year2005          -0.3028     0.6679  -0.45   0.651
```

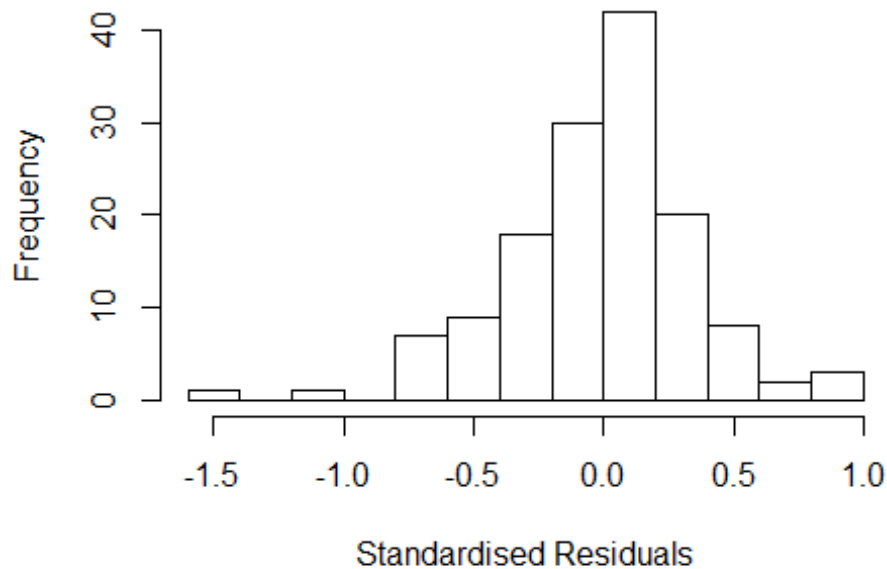


```

## Year2006          -0.3903      0.6854   -0.57    0.570
## Year2007          -0.4341      0.6680   -0.65    0.517
## Year2008          -0.4419      0.6693   -0.66    0.510
## Year2009          -0.4350      0.6888   -0.63    0.529
## Year2010          -0.2782      0.6980   -0.40    0.691
## Year2011          -0.3373      0.6730   -0.50    0.617
## Year2012          -0.4784      0.6661   -0.72    0.474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.171, Adjusted R-squared:  0.0486
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 124 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0138 0.8790 0.9510 0.8910 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      7.09e-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 4.517 1      2.125
## Year              4.517 16      1.048

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.4176 -0.2011 0.0223 0.1963 0.9924
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4176 0.7989 1.77 0.078 .
## FirstAuthorFemale1 0.0926 0.0805 1.15 0.252
## Year1997 -0.0895 0.8014 -0.11 0.911
## Year1998 0.1199 0.8158 0.15 0.883
## Year1999 -0.3000 0.8104 -0.37 0.712
## Year2000 -0.6314 0.7858 -0.80 0.423
## Year2001 -0.4147 0.7750 -0.54 0.594
## Year2002 -0.3279 0.7914 -0.41 0.679
## Year2003 -0.3129 0.7919 -0.40 0.693
## Year2004 -0.2283 0.7938 -0.29 0.774
## Year2005 -0.2132 0.7911 -0.27 0.788
## Year2006 -0.3536 0.8070 -0.44 0.662
```

```

## Year2007          -0.3691      0.7927   -0.47    0.642
## Year2008          -0.4130      0.7936   -0.52    0.604
## Year2009          -0.3730      0.8124   -0.46    0.647
## Year2010          -0.2469      0.8233   -0.30    0.765
## Year2011          -0.3006      0.8019   -0.37    0.708
## Year2012          -0.4481      0.7913   -0.57    0.572
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.331
## Multiple R-squared:  0.157, Adjusted R-squared:  0.0407
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0267 0.8720 0.9580 0.8930 0.9840 0.9980
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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 5.845 1          2.418
## Year          5.845 16          1.057

## [1] "List of 0 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
## -1.4952 -0.2165  0.0185  0.2025  0.9758

```

```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.4952    0.3661   4.08 7.9e-05 ***
## LastAuthorFemale1  0.1182    0.0912   1.30  0.197
## Year1997        -0.1471    0.3795  -0.39  0.699
## Year1998         0.0495    0.4104   0.12  0.904
## Year1999        -0.3728    0.3953  -0.94  0.348
## Year2000        -0.6421    0.3843  -1.67  0.097 .
## Year2001        -0.5179    0.3748  -1.38  0.170
## Year2002        -0.3354    0.3977  -0.84  0.401
## Year2003        -0.3966    0.3757  -1.06  0.293
## Year2004        -0.2839    0.3860  -0.74  0.463
## Year2005        -0.3204    0.3784  -0.85  0.399
## Year2006        -0.4423    0.3917  -1.13  0.261
## Year2007        -0.4679    0.3785  -1.24  0.219
## Year2008        -0.4718    0.3793  -1.24  0.216
## Year2009        -0.4912    0.4027  -1.22  0.225
## Year2010        -0.3084    0.4187  -0.74  0.463
## Year2011        -0.3919    0.3721  -1.05  0.294
## Year2012        -0.5096    0.3699  -1.38  0.171
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.165, Adjusted R-squared:  0.05
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| <= 0.00029 ( < 0.00071);
## 17 weights are ~= 1. The remaining 123 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.167  0.885   0.950   0.894   0.982   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          7.09e-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: 141"

```

```

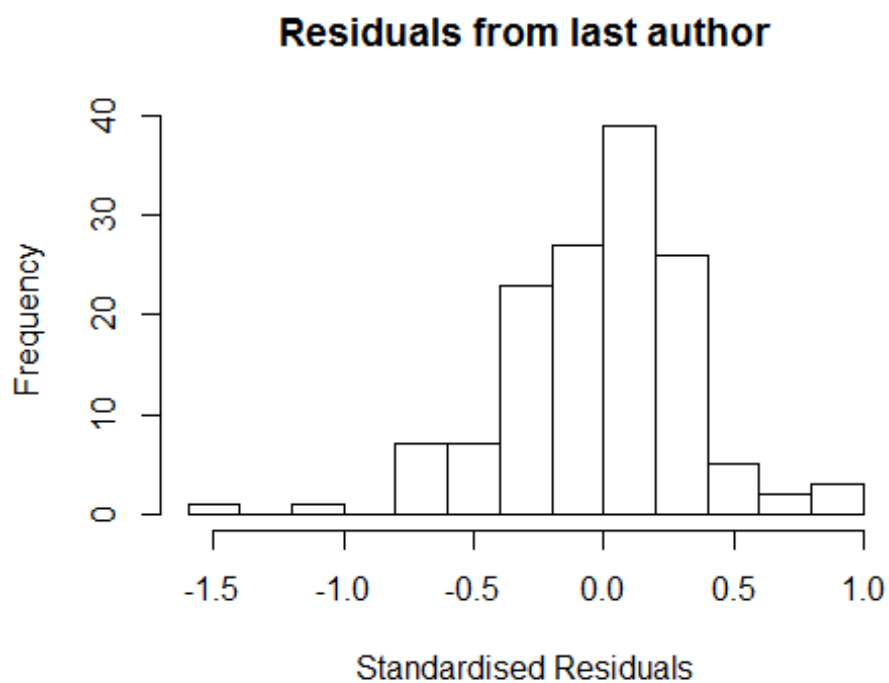
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2311"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    8   12   20   14   12    5   15   13   11   14   24   15   11    9
## 2011 2012
##   21   19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    2    4    4    3    4    1    4    6    9    5   14    9    6    4
## 2011 2012
##   11   10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    3    3    3    3    1    3    4    7    3   11    8    5    4
## 2011 2012
##   10    9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.608273453379"
## [1] "Male first author team size 2018 geometric mean: 3.28081836475037"

## 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 = 58, p-value = 0.8
## 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: 3.43697676676045"

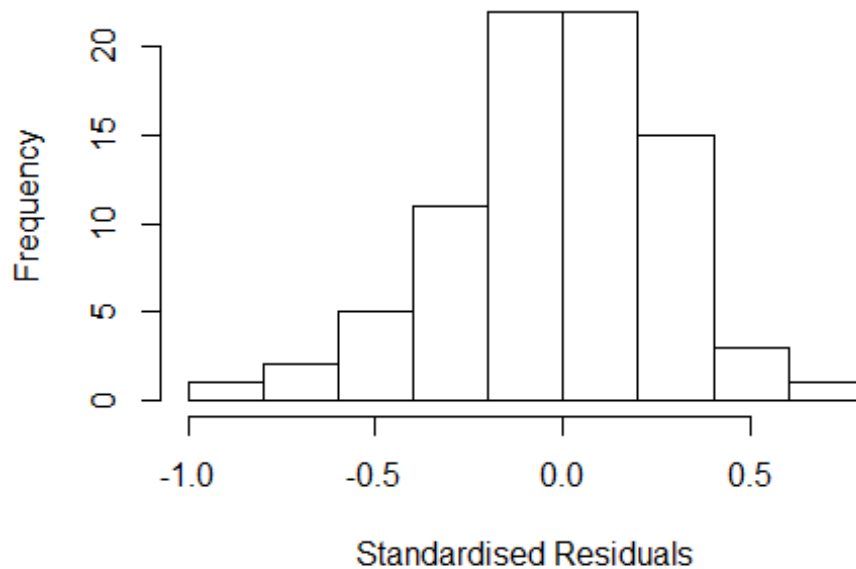
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 25, 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    3.405  1      1.845
## LastAuthorFemale     3.536  1      1.880
## UniqueAuthors    1755.165  4      2.544
## Year           7338.587 16      1.321
```

Residuals from first and last author and team size



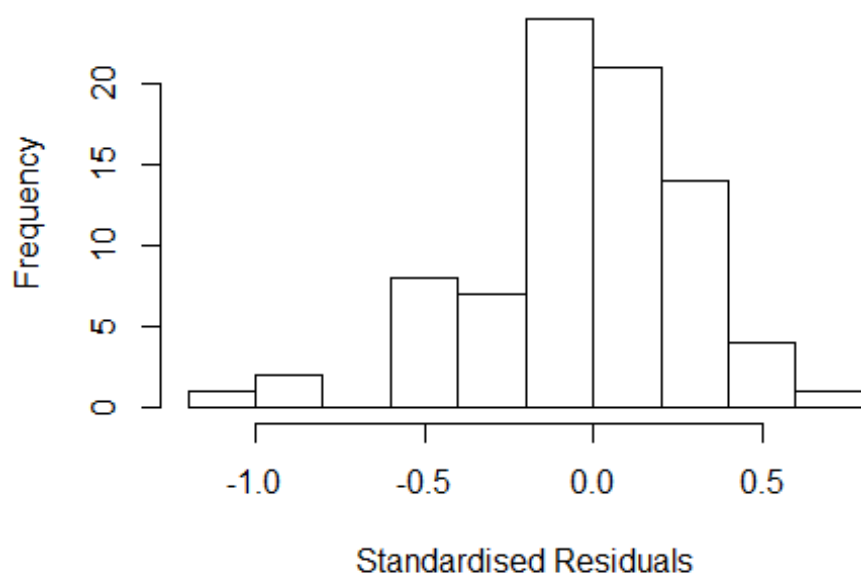
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.97878 -0.16918 0.00813 0.18927 0.68623
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.83242 0.36944 2.25 0.028 *
## FirstAuthorFemale1 -0.00567 0.09017 -0.06 0.950
## LastAuthorFemale1 0.26985 0.11385 2.37 0.021 *
## UniqueAuthors2 -0.11029 0.20080 -0.55 0.585
## UniqueAuthors3 0.08926 0.17820 0.50 0.618
## UniqueAuthors4 0.04019 0.19644 0.20 0.839
## UniqueAuthors5 0.20460 0.19908 1.03 0.308
## Year1997 0.62510 0.47820 1.31 0.196
## Year1998 0.66384 0.53597 1.24 0.220
## Year1999 0.58596 0.56282 1.04 0.302
```

```

## Year2000      0.21777      0.47889      0.45      0.651
## Year2001      0.18218      0.40056      0.45      0.651
## Year2002      0.88187      0.48446      1.82      0.074 .
## Year2003      0.07147      0.40616      0.18      0.861
## Year2004      0.18130      0.51663      0.35      0.727
## Year2005      0.23341      0.44786      0.52      0.604
## Year2006      0.14628      0.46250      0.32      0.753
## Year2007     -0.08059      0.47388     -0.17      0.866
## Year2008      0.14637      0.51447      0.28      0.777
## Year2009     -0.36465      0.45871     -0.79      0.430
## Year2010     -0.42894      0.48555     -0.88      0.381
## Year2011      0.23337      0.47277      0.49      0.623
## Year2012     -0.05428      0.45253     -0.12      0.905
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.321
## Multiple R-squared:  0.511, Adjusted R-squared:  0.328
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 75 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.333  0.889  0.964  0.920  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      1.22e-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 6.806 1 2.609
## LastAuthorFemale 2.356 1 1.535
## Year 14.610 16 1.087

```


Residuals from first and last author



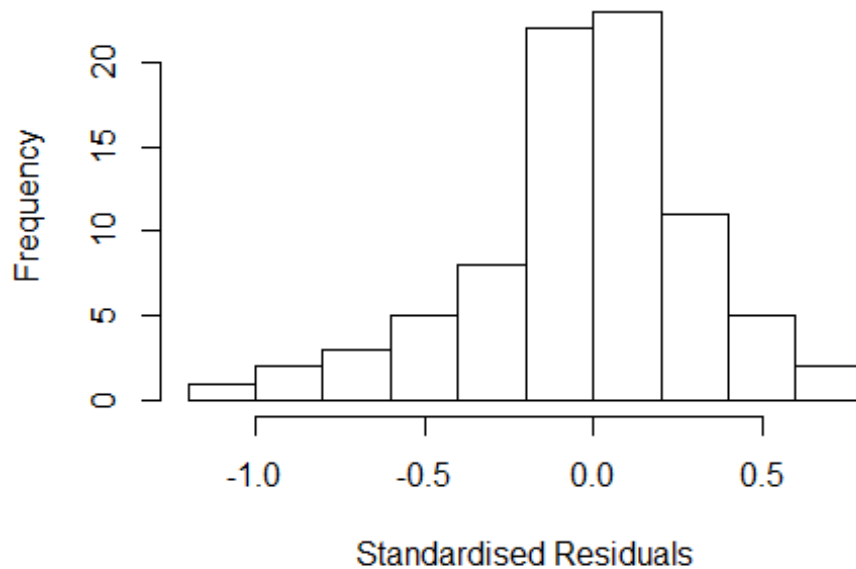
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.11671 -0.17004 -0.00164  0.18697  0.74393
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.82694    0.40876   2.02   0.0473 *
## FirstAuthorFemale1 -0.00709    0.10298  -0.07   0.9453
## LastAuthorFemale1  0.26078    0.09790   2.66   0.0098 **
## Year1997         0.62006    0.40969   1.51   0.1352
## Year1998         0.76009    0.49978   1.52   0.1333
## Year1999         0.60015    0.46672   1.29   0.2032
## Year2000         0.27144    0.39416   0.69   0.4936
## Year2001         0.21874    0.41329   0.53   0.5985
## Year2002         0.77706    0.40876   1.90   0.0619 .
## Year2003         0.12891    0.46455   0.28   0.7823
## Year2004         0.10967    0.46067   0.24   0.8126
## Year2005         0.24888    0.40533   0.61   0.5414
```

```

## Year2006          0.26262      0.41268      0.64      0.5268
## Year2007          0.05839      0.40315      0.14      0.8853
## Year2008          0.28977      0.40513      0.72      0.4771
## Year2009         -0.41687      0.45385     -0.92      0.3618
## Year2010         -0.34954      0.42807     -0.82      0.4173
## Year2011          0.29374      0.40825      0.72      0.4745
## Year2012          0.04360      0.40355      0.11      0.9143
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.303
## Multiple R-squared:  0.483, Adjusted R-squared:  0.335
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 78 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.145  0.890  0.963  0.906  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      1.22e-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 5.558 1      2.358
## Year              5.558 16      1.055

```

Residuals from first author



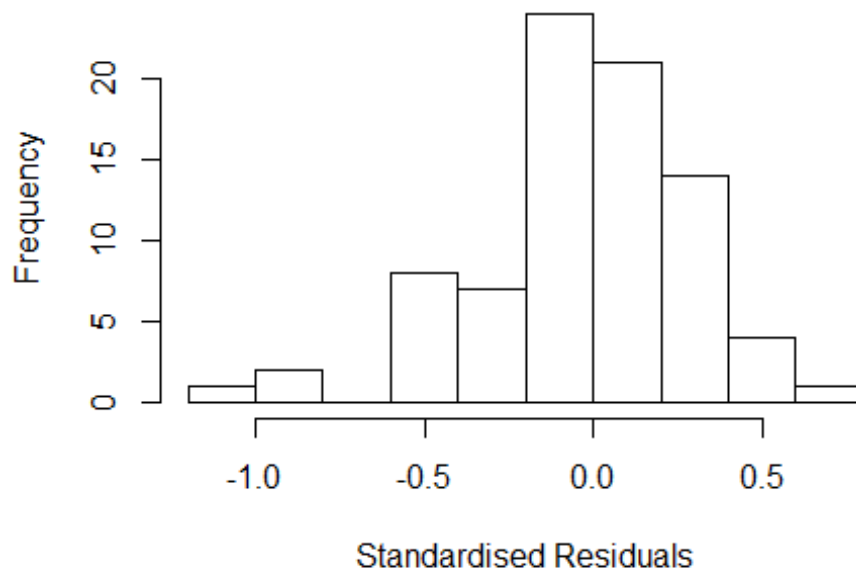
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.16740 -0.18083  0.00787  0.15351  0.69201
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.8000     0.3851   2.08  0.042 *
## FirstAuthorFemale1 0.0617     0.1113   0.55  0.582
## Year1997          0.6470     0.3861   1.68  0.099 .
## Year1998          0.7926     0.5140   1.54  0.128
## Year1999          0.6551     0.5428   1.21  0.232
## Year2000          0.2522     0.3667   0.69  0.494
## Year2001          0.3116     0.3791   0.82  0.414
## Year2002          0.8040     0.3851   2.09  0.041 *
## Year2003          0.1392     0.4270   0.33  0.745
## Year2004          0.1366     0.4428   0.31  0.759
## Year2005          0.3617     0.3816   0.95  0.347
## Year2006          0.2895     0.3893   0.74  0.460
```

```

## Year2007          0.1616      0.3708      0.44      0.664
## Year2008          0.3674      0.3791      0.97      0.336
## Year2009         -0.3381      0.4548     -0.74      0.460
## Year2010         -0.3452      0.4034     -0.86      0.395
## Year2011          0.3565      0.3903      0.91      0.364
## Year2012          0.0966      0.3806      0.25      0.801
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.295
## Multiple R-squared:  0.44,   Adjusted R-squared:  0.292
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0829 0.8520 0.9570 0.8860 0.9910 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      1.22e-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 2.344 1      1.531
## Year            2.344 16      1.027

```

Residuals from last author



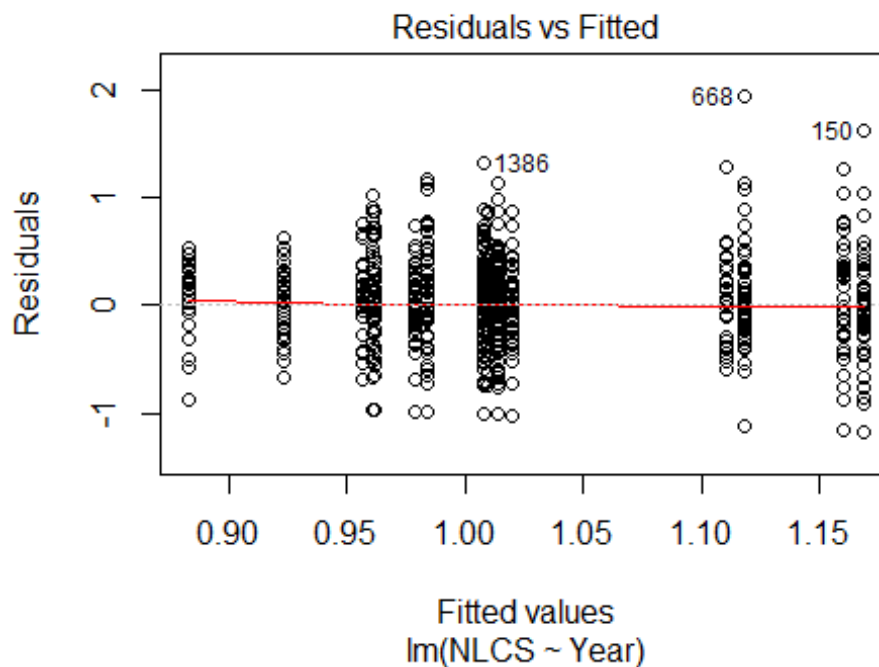
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11800 -0.17106 -0.00282 0.18196 0.74739
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8311 0.3924 2.12 0.038 *
## LastAuthorFemale1 0.2576 0.0990 2.60 0.011 *
## Year1997 0.6159 0.3934 1.57 0.122
## Year1998 0.7579 0.4807 1.58 0.120
## Year1999 0.5962 0.4535 1.31 0.193
## Year2000 0.2625 0.3942 0.67 0.508
## Year2001 0.2131 0.4051 0.53 0.601
## Year2002 0.7729 0.3924 1.97 0.053 .
## Year2003 0.1224 0.4517 0.27 0.787
## Year2004 0.1055 0.4479 0.24 0.815
## Year2005 0.2432 0.3978 0.61 0.543
## Year2006 0.2584 0.3965 0.65 0.517
```

```

## Year2007          0.0553      0.3876      0.14      0.887
## Year2008          0.2869      0.3978      0.72      0.473
## Year2009         -0.4245      0.4471     -0.95      0.346
## Year2010         -0.3557      0.4258     -0.84      0.407
## Year2011          0.2896      0.3935      0.74      0.464
## Year2012          0.0374      0.4000      0.09      0.926
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.299
## Multiple R-squared:  0.486, Adjusted R-squared:  0.35
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 78 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.130  0.885   0.964   0.903   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      1.22e-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: 82"
## [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
##   64   91   71  111   80  110   93   92   60   63   68   68   72   65   61
## 2011 2012
##  100   88
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   34   48   42   56   34   58   59   57   37   47   47   30   43   47   37
## 2011 2012

```

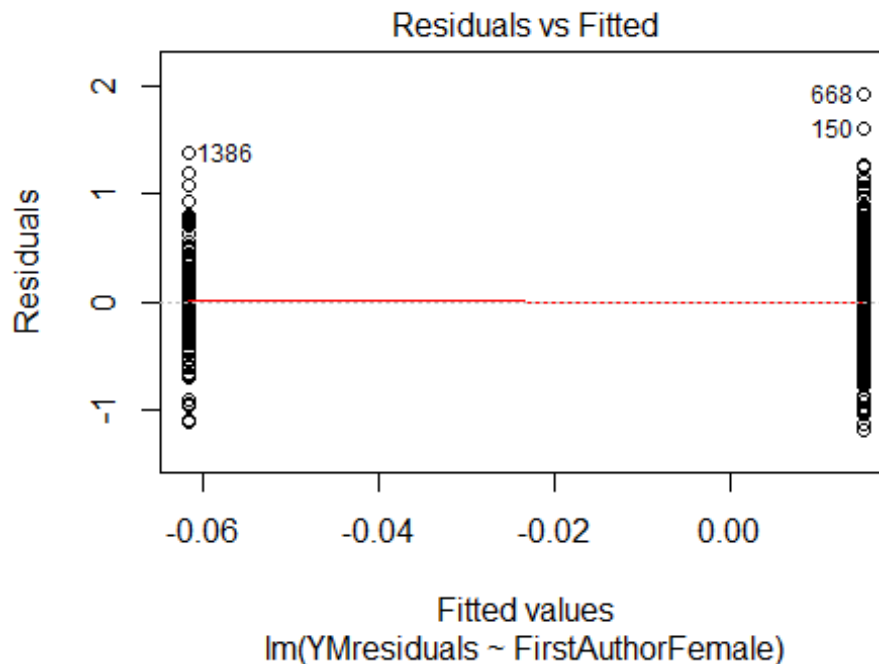
```
## 47 43
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 29 42 40 50 27 56 56 52 35 41 41 25 34 43 35
## 2011 2012
## 41 42
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 0.007, df = 1, p-value = 0.9
## [1] "Female first author team size 2018 geometric mean: 3.81501248899068"
## [1] "Male first author team size 2018 geometric mean: 3.308974752555"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

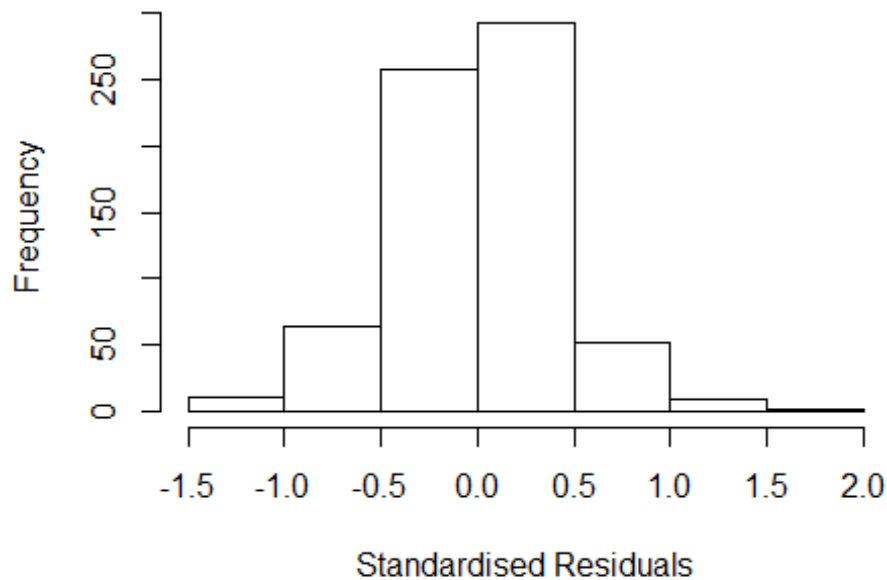
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 450, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.9328385885927"
## [1] "Male last author team size 2018 geometric mean: 3.64830984674394"

## 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 = 220, 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 1.157  1      1.076
## LastAuthorFemale  1.433  1      1.197
## UniqueAuthors    1.936  4      1.086
## Year              2.622 16      1.031
```


Residuals from first and last author and team size



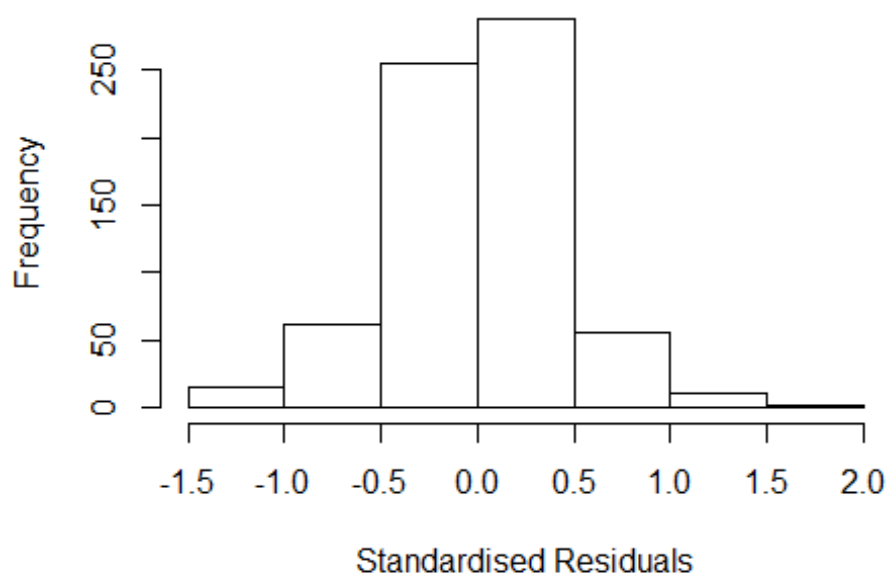
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2038 -0.2513  0.0155  0.2512  1.9658
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0042    0.0836   12.01  < 2e-16 ***
## FirstAuthorFemale1 -0.1098    0.0397   -2.77  0.0058 **
## LastAuthorFemale1  0.0459    0.0411    1.12  0.2643
## UniqueAuthors2     0.0726    0.0468    1.55  0.1215
## UniqueAuthors3     0.1271    0.0510    2.49  0.0129 *
## UniqueAuthors4     0.1023    0.0554    1.85  0.0652 .
## UniqueAuthors5     0.2432    0.0577    4.22 2.8e-05 ***
## Year1997           0.0582    0.1152    0.51  0.6132
## Year1998           0.0527    0.1207    0.44  0.6624
## Year1999          -0.0946    0.1180   -0.80  0.4229
```

```

## Year2000          -0.2043      0.1118    -1.83    0.0682 .
## Year2001          -0.0650      0.0969    -0.67    0.5028
## Year2002           0.0154      0.0971     0.16    0.8743
## Year2003          -0.0598      0.0917    -0.65    0.5148
## Year2004          -0.0812      0.1131    -0.72    0.4730
## Year2005          -0.0384      0.0975    -0.39    0.6938
## Year2006          -0.1463      0.0967    -1.51    0.1306
## Year2007          -0.1969      0.1058    -1.86    0.0633 .
## Year2008          -0.0831      0.0968    -0.86    0.3910
## Year2009          -0.0849      0.1057    -0.80    0.4222
## Year2010          -0.0896      0.1409    -0.64    0.5248
## Year2011          -0.0360      0.1091    -0.33    0.7413
## Year2012          -0.0490      0.1012    -0.48    0.6286
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.377
## Multiple R-squared:  0.0616, Adjusted R-squared:  0.0306
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 261 is an outlier with |weight| = 0 ( < 0.00015);
## 69 weights are ~= 1. The remaining 619 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0013 0.8610 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          1.45e-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 1.114 1 1.055
## LastAuthorFemale 1.379 1 1.174
## Year 1.435 16 1.011

```

Residuals from first and last author



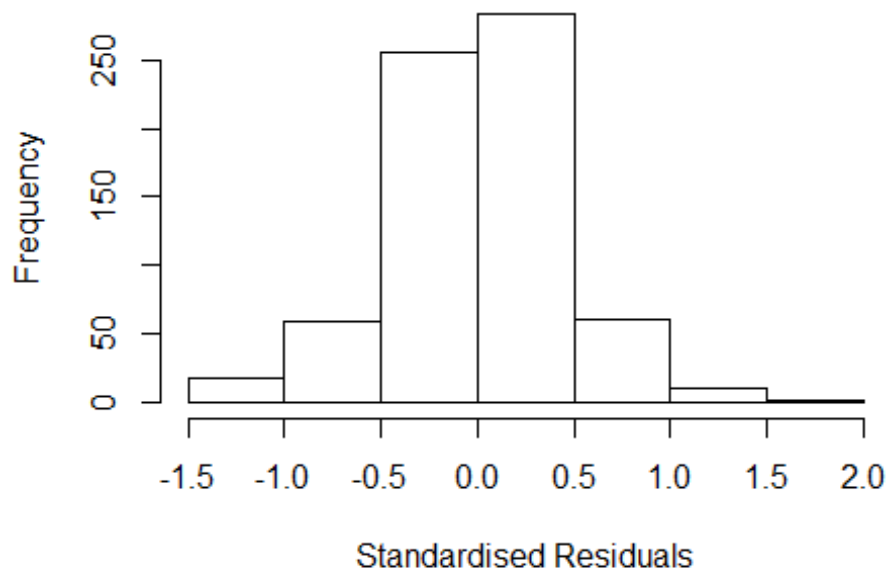
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1431 -0.2495 0.0142 0.2537 1.9713
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06212 0.07769 13.67 <2e-16 ***
## FirstAuthorFemale1 -0.10480 0.04014 -2.61 0.0092 **
## LastAuthorFemale1 0.05659 0.04099 1.38 0.1678
## Year1997 0.08102 0.11042 0.73 0.4633
## Year1998 0.06916 0.11702 0.59 0.5547
## Year1999 -0.09302 0.11487 -0.81 0.4183
## Year2000 -0.19869 0.11329 -1.75 0.0799 .
## Year2001 -0.05557 0.09505 -0.58 0.5590
## Year2002 0.02456 0.09351 0.26 0.7930
## Year2003 -0.05490 0.08981 -0.61 0.5412
## Year2004 -0.06598 0.11323 -0.58 0.5603
## Year2005 -0.00954 0.09416 -0.10 0.9193
```

```

## Year2006          -0.10097    0.09290   -1.09    0.2775
## Year2007          -0.12607    0.10136   -1.24    0.2140
## Year2008          -0.04230    0.09451   -0.45    0.6546
## Year2009          -0.07590    0.10068   -0.75    0.4512
## Year2010          -0.04840    0.13650   -0.35    0.7230
## Year2011           0.01068    0.10756    0.10    0.9209
## Year2012           0.00657    0.09759    0.07    0.9463
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.0343, Adjusted R-squared:  0.00838
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 261 is an outlier with |weight| = 0 ( < 0.00015);
## 57 weights are ~= 1. The remaining 631 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0288 0.8640 0.9540 0.8900 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      1.45e-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 1.059 1          1.029
## Year              1.059 16          1.002

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.1660 -0.2569  0.0122  0.2572  1.9698
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0658     0.0765   13.93  <2e-16 ***
## FirstAuthorFemale1 -0.0900     0.0393   -2.29   0.022 *
## Year1997          0.1001     0.1067    0.94   0.348
## Year1998          0.0759     0.1161    0.65   0.513
## Year1999         -0.0890     0.1152   -0.77   0.440
## Year2000         -0.1919     0.1130   -1.70   0.090 .
## Year2001         -0.0519     0.0936   -0.55   0.579
## Year2002          0.0224     0.0926    0.24   0.809
## Year2003         -0.0488     0.0883   -0.55   0.580
## Year2004         -0.0593     0.1117   -0.53   0.596
## Year2005         -0.0053     0.0928   -0.06   0.954
## Year2006         -0.1026     0.0919   -1.12   0.265
```

```

## Year2007          -0.1139      0.0997   -1.14    0.254
## Year2008          -0.0339      0.0935   -0.36    0.717
## Year2009          -0.0660      0.0990   -0.67    0.505
## Year2010          -0.0519      0.1348   -0.38    0.701
## Year2011           0.0152      0.1061    0.14    0.886
## Year2012           0.0105      0.0965    0.11    0.913
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.0323, Adjusted R-squared:  0.00783
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 261 is an outlier with |weight| = 0 ( < 0.00015);
## 59 weights are ~= 1. The remaining 629 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0368 0.8620 0.9520 0.8900 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          1.45e-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 1.314 1          1.146
## Year              1.314 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max

```

```

## -1.149 -0.256 0.015 0.264 1.987
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.049298   0.080370   13.06  <2e-16 ***
## LastAuthorFemale1 0.029290   0.040320    0.73   0.468
## Year1997        0.098900   0.113022    0.88   0.382
## Year1998        0.070829   0.121274    0.58   0.559
## Year1999       -0.101731   0.117031   -0.87   0.385
## Year2000       -0.194301   0.116263   -1.67   0.095 .
## Year2001       -0.057414   0.098034   -0.59   0.558
## Year2002        0.021892   0.096116    0.23   0.820
## Year2003       -0.053774   0.092133   -0.58   0.560
## Year2004       -0.072315   0.113606   -0.64   0.525
## Year2005       -0.011289   0.096869   -0.12   0.907
## Year2006       -0.104448   0.096052   -1.09   0.277
## Year2007       -0.117389   0.103708   -1.13   0.258
## Year2008       -0.044646   0.098133   -0.45   0.649
## Year2009       -0.081063   0.101996   -0.79   0.427
## Year2010       -0.063452   0.138924   -0.46   0.648
## Year2011       -0.000796   0.108661   -0.01   0.994
## Year2012       -0.006709   0.100543   -0.07   0.947
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.0262, Adjusted R-squared:  0.00148
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 261 is an outlier with |weight| = 0 ( < 0.00015);
## 64 weights are ~= 1. The remaining 624 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0302 0.8500 0.9510 0.8880 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           1.45e-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)

```

```

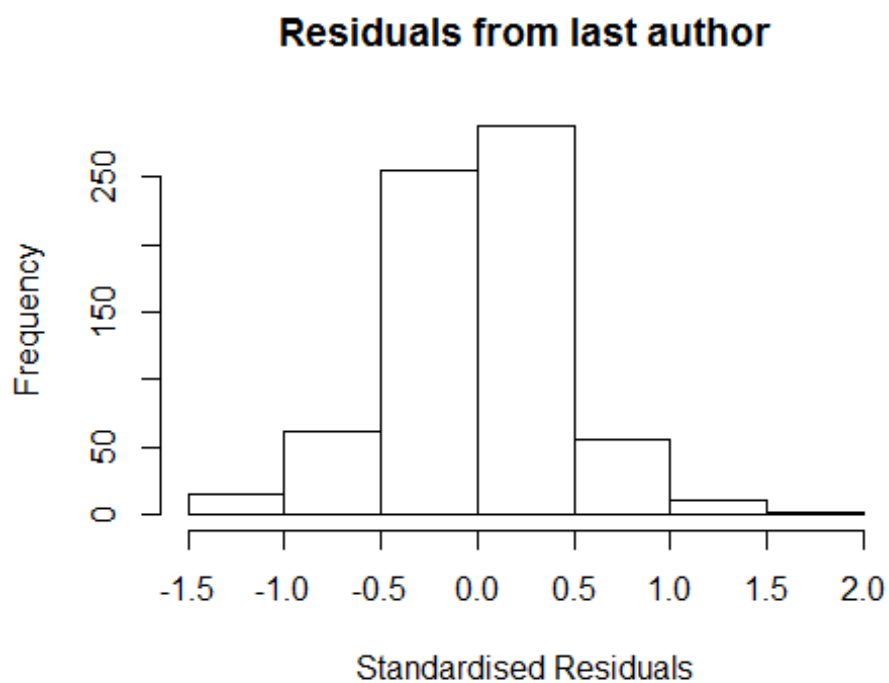
## [1] "Sample size for the above analysis: 689"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2400"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    8    2    8    5    7    9    8    9    8    9   10   16    7
## 2011 2012
##   21   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    5    1    3    1    4    7    7    5    7    6    8    8    2
## 2011 2012
##   14    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    4    1    3    0    3    7    6    5    6    5    7    6    2
## 2011 2012
##   13    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.09062348923505"
## [1] "Male first author team size 2018 geometric mean: 3.53602086097445"

## 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 = 17, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.52701905583787"
## [1] "Male last author team size 2018 geometric mean: 3.33702088205365"

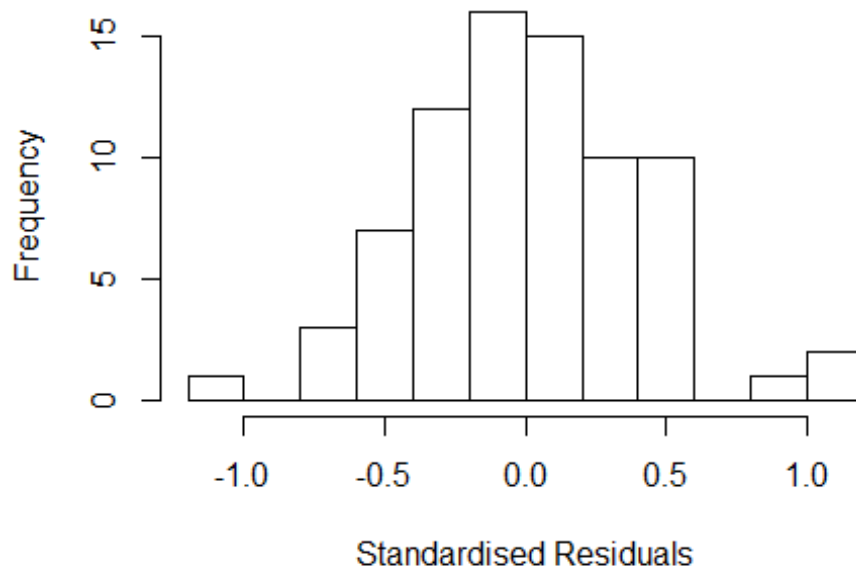
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20, 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 3.273e+00  1      1.809e+00
## LastAuthorFemale  3.117e+13  1      5.583e+06
## UniqueAuthors    2.535e+02  4      1.998e+00
## Year              4.602e+15 15      3.327e+00
```

Residuals from first and last author and team size



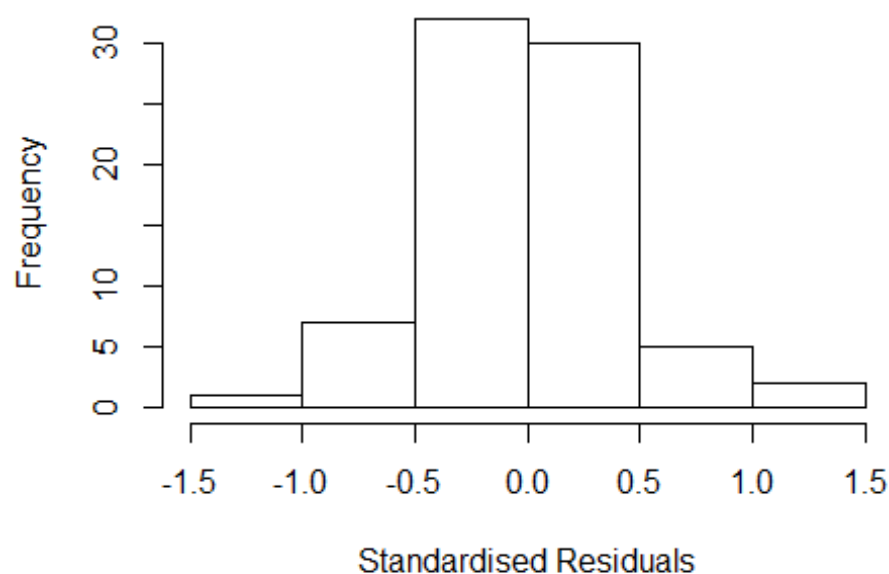
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.00e+00 -2.68e-01 -3.33e-16  2.96e-01  1.16e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.1199    0.1769    0.68  0.50090
## FirstAuthorFemale1 0.2014    0.1270    1.59  0.11848
## LastAuthorFemale1 0.0733    0.2010    0.36  0.71689
## UniqueAuthors2    0.1146    0.2182    0.53  0.60147
## UniqueAuthors3   -0.1903    0.2632   -0.72  0.47270
## UniqueAuthors4   -0.2073    0.1822   -1.14  0.26016
## UniqueAuthors5   -0.1199    0.1769   -0.68  0.50090
## Year1997         1.8058    0.5076    3.56  0.00078 ***
## Year1998         1.5278    0.1991    7.67  3.0e-10 ***
## Year1999         0.7457    0.2010    3.71  0.00048 ***
```

```

## Year2000          1.4979      0.2917      5.14  3.8e-06 ***
## Year2002          1.3563      0.2620      5.18  3.3e-06 ***
## Year2003          1.2705      0.3257      3.90  0.00026 ***
## Year2004          1.2608      0.2623      4.81  1.2e-05 ***
## Year2005          0.9480      0.2218      4.27  7.7e-05 ***
## Year2006          1.1095      0.2895      3.83  0.00033 ***
## Year2007          1.2723      0.2522      5.04  5.3e-06 ***
## Year2008          1.1021      0.1528      7.21  1.7e-09 ***
## Year2009          1.4717      0.2376      6.19  7.8e-08 ***
## Year2010          0.4163      0.6373      0.65  0.51627
## Year2011          1.0272      0.1774      5.79  3.5e-07 ***
## Year2012          0.7968      0.4893      1.63  0.10917
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.393, Adjusted R-squared:  0.161
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.348  0.871  0.930  0.898  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      1.30e-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 2.072e+00 1      1.44
## LastAuthorFemale -5.629e+14 1      NaN
## Year              -1.118e+15 15      NaN

```

Residuals from first and last author



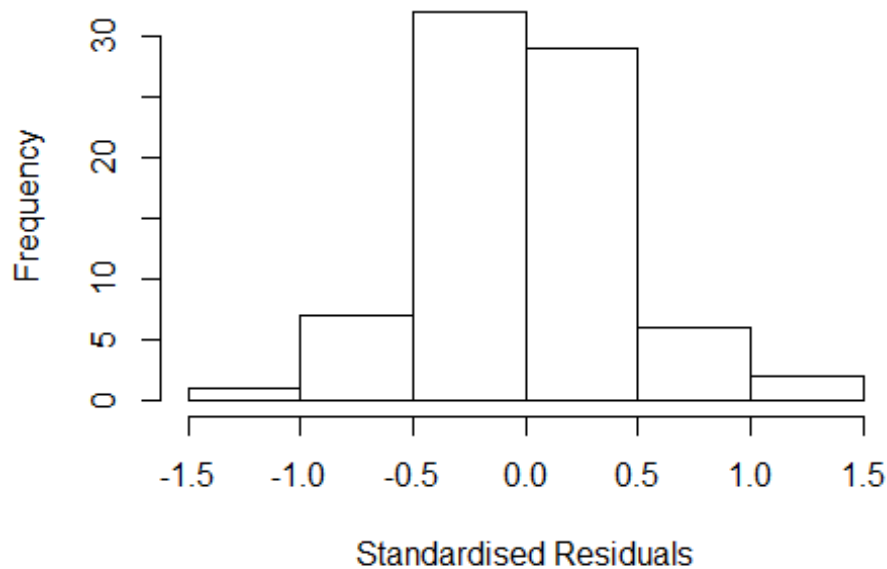
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.13e+00 -2.51e-01 -8.88e-16 3.06e-01 1.21e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.53e-16 1.63e-08 0.00 1.00000
## FirstAuthorFemale1 1.94e-01 1.18e-01 1.64 0.10718
## LastAuthorFemale1 8.26e-02 1.79e-01 0.46 0.64566
## Year1997 1.98e+00 3.68e-01 5.38 1.3e-06 ***
## Year1998 1.58e+00 2.13e-01 7.46 4.6e-10 ***
## Year1999 7.36e-01 1.79e-01 4.12 0.00012 ***
## Year2000 1.56e+00 2.95e-01 5.29 1.9e-06 ***
## Year2002 1.38e+00 2.18e-01 6.33 3.6e-08 ***
## Year2003 1.36e+00 2.70e-01 5.05 4.6e-06 ***
## Year2004 1.30e+00 2.13e-01 6.12 8.1e-08 ***
## Year2005 9.50e-01 2.34e-01 4.06 0.00015 ***
## Year2006 1.18e+00 1.56e-01 7.56 3.1e-10 ***
```

```

## Year2007          1.28e+00  1.90e-01  6.74  7.5e-09 ***
## Year2008          1.07e+00  1.10e-01  9.74  6.7e-14 ***
## Year2009          1.45e+00  1.42e-01  10.22  1.1e-14 ***
## Year2010          5.94e-01  7.14e-01  0.83  0.40944
## Year2011          1.11e+00  9.73e-02  11.37  < 2e-16 ***
## Year2012          8.58e-01  3.22e-01  2.66  0.01002 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.341, Adjusted R-squared:  0.152
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.373  0.892  0.953  0.907  0.984  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 2.016 1          1.420
## Year              2.016 15          1.024

```

Residuals from first author



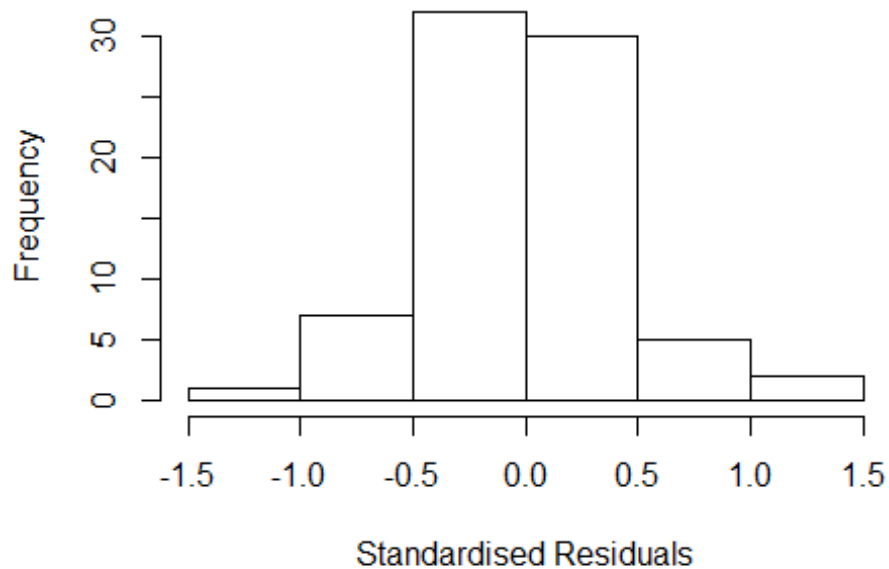
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.04e+00 -2.66e-01 2.22e-16 3.06e-01 1.24e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.82e-15 3.75e-08 0.00e+00 1.0000
## FirstAuthorFemale1 2.14e-01 1.25e-01 1.71e+00 0.0929 .
## Year1997 1.98e+00 3.54e-01 5.60e+00 5.6e-07 ***
## Year1998 1.59e+00 2.08e-01 7.65e+00 2.0e-10 ***
## Year1999 8.19e-01 1.74e-08 4.72e+07 < 2e-16 ***
## Year2000 1.58e+00 3.06e-01 5.16e+00 2.9e-06 ***
## Year2002 1.38e+00 2.14e-01 6.45e+00 2.2e-08 ***
## Year2003 1.39e+00 2.38e-01 5.85e+00 2.2e-07 ***
## Year2004 1.30e+00 1.98e-01 6.60e+00 1.2e-08 ***
## Year2005 9.47e-01 2.17e-01 4.36e+00 5.1e-05 ***
## Year2006 1.21e+00 1.68e-01 7.18e+00 1.2e-09 ***
## Year2007 1.29e+00 1.82e-01 7.09e+00 1.8e-09 ***
```

```

## Year2008          1.09e+00  9.91e-02 1.10e+01  4.8e-16 ***
## Year2009          1.46e+00  1.31e-01 1.11e+01  3.2e-16 ***
## Year2010          5.93e-01  6.36e-01 9.30e-01   0.3546
## Year2011          1.11e+00  9.39e-02 1.18e+01  < 2e-16 ***
## Year2012          8.22e-01  2.89e-01 2.84e+00   0.0061 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.452
## Multiple R-squared:  0.332, Adjusted R-squared:  0.154
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.430  0.905  0.959  0.921  0.985  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      1.30e-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 2.896e+13  1      5.382e+06
## Year             2.896e+13 15      2.810e+00

```

Residuals from last author



```
## [1] "List of 0 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
## -1.098340 -0.238177 -0.000615 0.279207 1.370394
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) -1.52e-15 3.52e-08 0.00 1.00000
## LastAuthorFemale1 1.06e-01 1.78e-01 0.60 0.55158
## Year1997 1.98e+00 3.69e-01 5.37 1.3e-06 ***
## Year1998 1.58e+00 2.13e-01 7.44 4.4e-10 ***
## Year1999 7.13e-01 1.78e-01 4.01 0.00017 ***
## Year2000 1.60e+00 3.61e-01 4.45 3.8e-05 ***
## Year2002 1.38e+00 2.19e-01 6.33 3.5e-08 ***
## Year2003 1.35e+00 2.71e-01 5.00 5.3e-06 ***
## Year2004 1.38e+00 1.70e-01 8.11 3.2e-11 ***
## Year2005 1.02e+00 2.15e-01 4.72 1.5e-05 ***
## Year2006 1.19e+00 1.70e-01 7.00 2.5e-09 ***
## Year2007 1.32e+00 1.93e-01 6.83 4.8e-09 ***
```



```

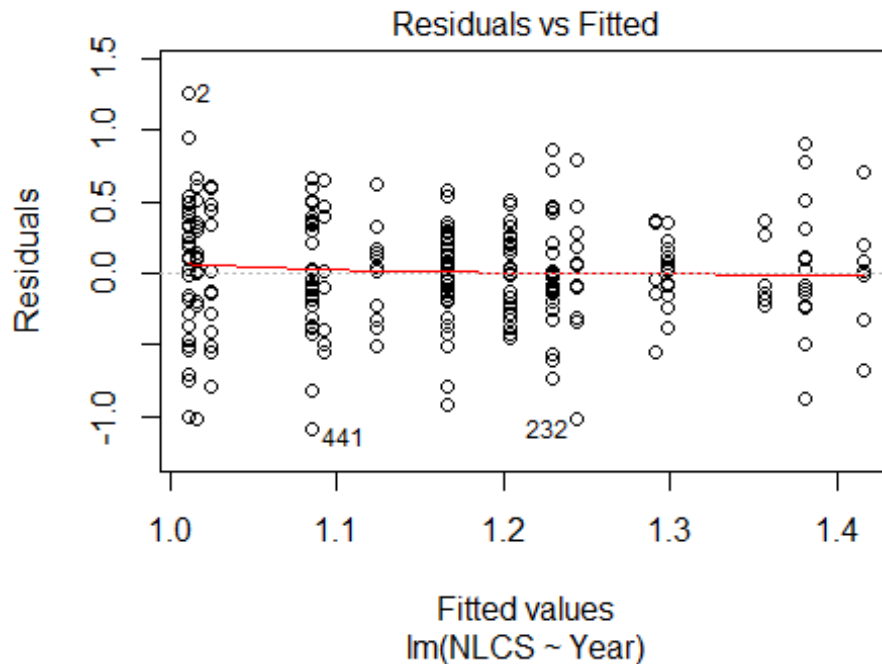
## Year2008          1.09e+00   9.83e-02   11.11   3.4e-16 ***
## Year2009          1.48e+00   1.40e-01   10.57   2.5e-15 ***
## Year2010          5.93e-01   7.20e-01    0.82   0.41304
## Year2011          1.19e+00   9.08e-02   13.08   < 2e-16 ***
## Year2012          9.92e-01   3.19e-01    3.11   0.00287 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.327, Adjusted R-squared:  0.148
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.247  0.908  0.954  0.902  0.977  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      1.30e-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: 77"
## [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]"
##
## 2012
##    1
##
## 2012
##    1
##
## 2012
##    1
## [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: 3"
## [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: 1"
## [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
## 38 30 28 26 21 30 26 23 24 25 27 39 30 41 38
## 2011 2012
## 26 29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 19 7 6 17 5 8 11 11 14 15 14 23 25 23 25
## 2011 2012
## 16 17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 19 7 5 15 4 8 10 6 12 9 13 20 18 18 20
## 2011 2012
## 14 16
## [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

```



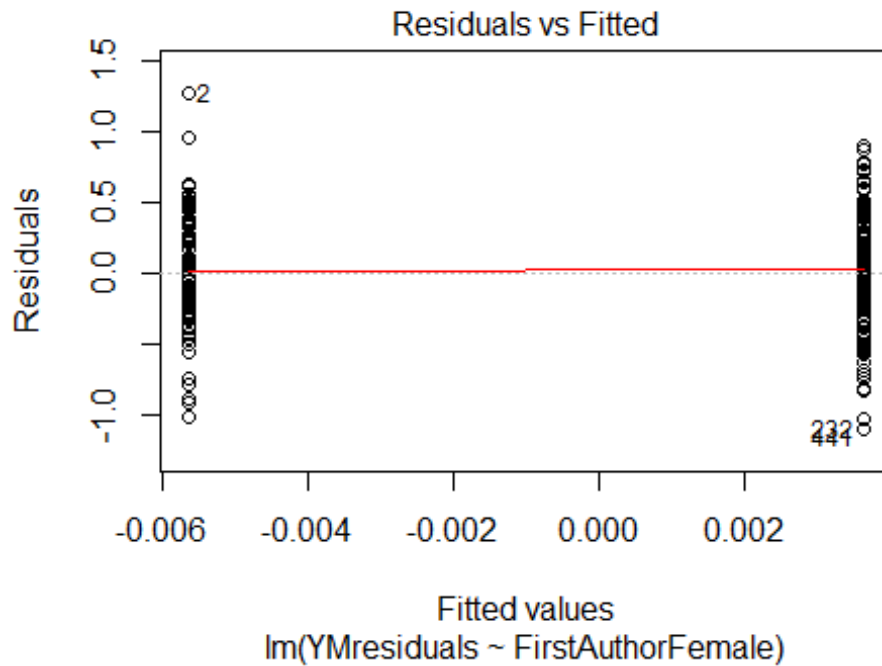
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.013, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 4.67788567485604"
## [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

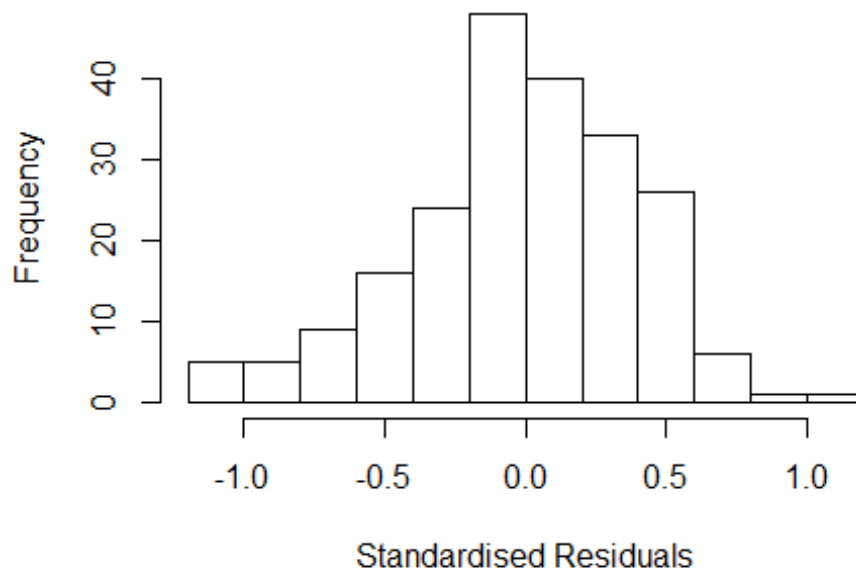
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8.5, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.47213595499958"
## [1] "Male last author team size 2018 geometric mean: 3.39043640614487"

## 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 = 6.5, 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 1.613 1      1.270
## LastAuthorFemale  1.580 1      1.257
## UniqueAuthors     5.299 4      1.232
## Year              9.809 16     1.074
```

Residuals from first and last author and team size



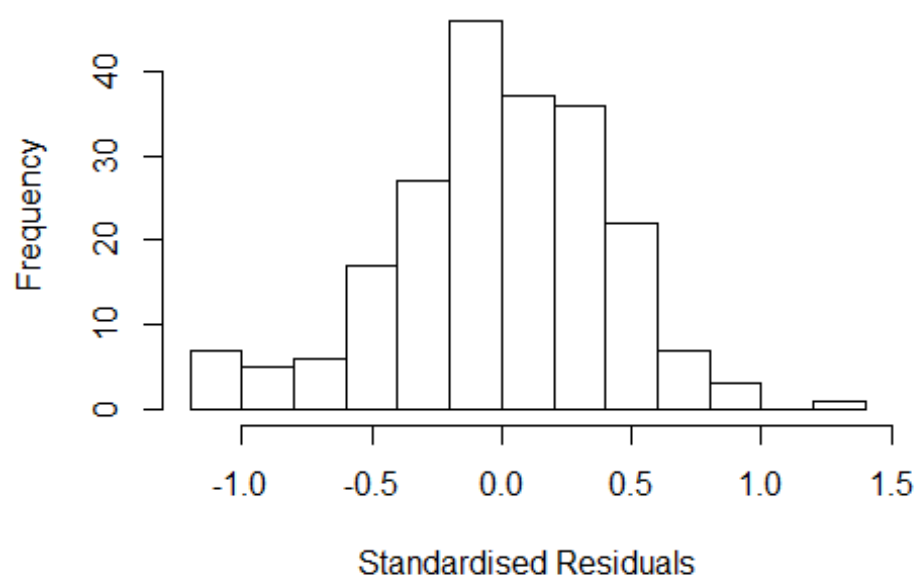
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.120586 -0.238380 0.000448 0.276867 1.180219
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7849 0.1664 4.72 4.6e-06 ***
## FirstAuthorFemale1 -0.0367 0.0644 -0.57 0.569
## LastAuthorFemale1 -0.0678 0.0876 -0.77 0.440
## UniqueAuthors2 0.2182 0.1489 1.47 0.144
## UniqueAuthors3 0.3406 0.1441 2.36 0.019 *
## UniqueAuthors4 0.2756 0.1509 1.83 0.069 .
## UniqueAuthors5 0.2769 0.1542 1.80 0.074 .
## Year1997 0.4184 0.2115 1.98 0.049 *
## Year1998 0.3376 0.1679 2.01 0.046 *
## Year1999 0.0747 0.1928 0.39 0.699
```

```

## Year2000          0.1630      0.2202      0.74      0.460
## Year2001          0.0686      0.2120      0.32      0.747
## Year2002          0.2481      0.1809      1.37      0.172
## Year2003          0.1903      0.2741      0.69      0.488
## Year2004          0.0620      0.2122      0.29      0.771
## Year2005          0.2591      0.1599      1.62      0.107
## Year2006          0.1630      0.1722      0.95      0.345
## Year2007          0.2051      0.1754      1.17      0.244
## Year2008          0.0308      0.1946      0.16      0.874
## Year2009          0.1169      0.1614      0.72      0.469
## Year2010          0.1279      0.1561      0.82      0.414
## Year2011          0.2612      0.1768      1.48      0.141
## Year2012          0.0955      0.1727      0.55      0.581
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.103, Adjusted R-squared:  0.00011
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 191 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.359  0.877  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      4.67e-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 1.563 1      1.250
## LastAuthorFemale  1.476 1      1.215
## Year              2.249 16      1.026

```

Residuals from first and last author



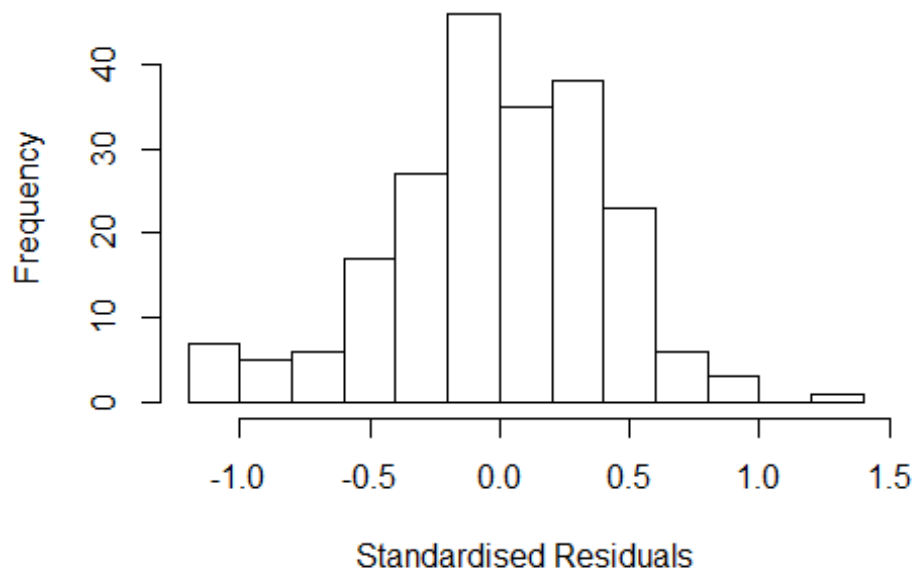
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11025 -0.24650 -0.00188 0.27155 1.28568
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98951 0.17244 5.74 3.6e-08 ***
## FirstAuthorFemale1 -0.00619 0.06502 -0.10 0.924
## LastAuthorFemale1 -0.03974 0.08739 -0.45 0.650
## Year1997 0.43183 0.22929 1.88 0.061 .
## Year1998 0.40992 0.20261 2.02 0.044 *
## Year1999 0.10638 0.22455 0.47 0.636
## Year2000 0.21368 0.24802 0.86 0.390
## Year2001 0.11607 0.25581 0.45 0.651
## Year2002 0.27003 0.22513 1.20 0.232
## Year2003 0.09087 0.24650 0.37 0.713
## Year2004 0.10569 0.24324 0.43 0.664
## Year2005 0.35400 0.18513 1.91 0.057 .
```

```

## Year2006          0.22758      0.20290      1.12      0.263
## Year2007          0.26067      0.20132      1.29      0.197
## Year2008          0.10915      0.21253      0.51      0.608
## Year2009          0.15674      0.18888      0.83      0.408
## Year2010          0.18339      0.18718      0.98      0.328
## Year2011          0.30079      0.20339      1.48      0.141
## Year2012          0.12694      0.20104      0.63      0.529
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.386
## Multiple R-squared:  0.0676, Adjusted R-squared:  -0.0184
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 185 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.243  0.864  0.939  0.886  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      4.67e-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 1.501 1      1.225
## Year              1.501 16      1.013

```


Residuals from first author



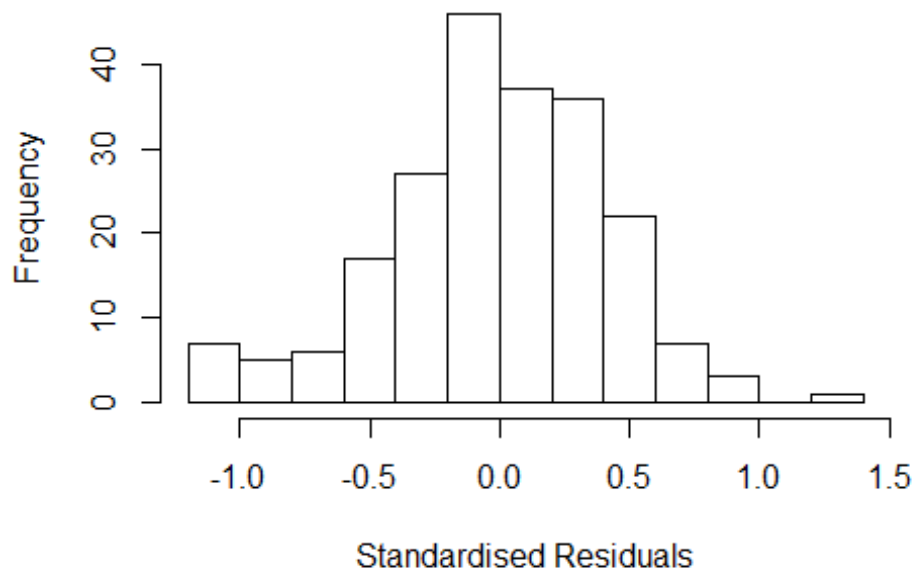
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.098543 -0.243803 -0.000604 0.265450 1.296748
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.97523 0.16749 5.82 2.3e-08 ***
## FirstAuthorFemale1 -0.00298 0.06399 -0.05 0.963
## Year1997 0.43846 0.22859 1.92 0.057 .
## Year1998 0.42424 0.19846 2.14 0.034 *
## Year1999 0.10396 0.22062 0.47 0.638
## Year2000 0.22698 0.24458 0.93 0.355
## Year2001 0.10938 0.25529 0.43 0.669
## Year2002 0.26577 0.22468 1.18 0.238
## Year2003 0.10477 0.24293 0.43 0.667
## Year2004 0.11923 0.23728 0.50 0.616
## Year2005 0.35897 0.18248 1.97 0.051 .
## Year2006 0.23645 0.19907 1.19 0.236
```

```

## Year2007          0.26929    0.19862    1.36    0.177
## Year2008          0.11471    0.20822    0.55    0.582
## Year2009          0.16727    0.18453    0.91    0.366
## Year2010          0.19007    0.18492    1.03    0.305
## Year2011          0.30025    0.20291    1.48    0.141
## Year2012          0.12629    0.19869    0.64    0.526
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.0662, Adjusted R-squared:  -0.0148
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 184 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.242  0.865   0.943   0.887   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      4.67e-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 1.441 1          1.200
## Year              1.441 16          1.011

```

Residuals from last author



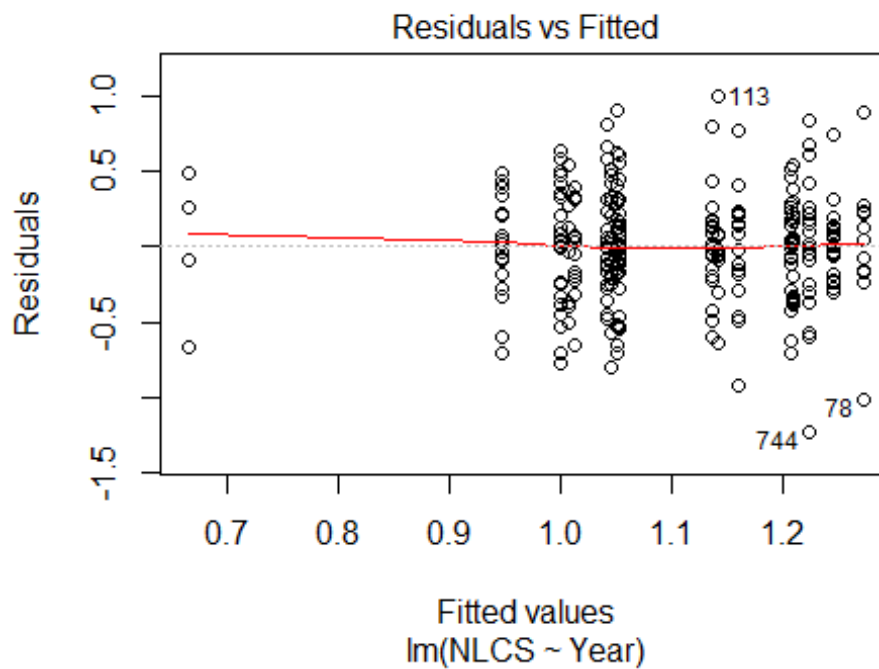
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.111814 -0.249230 -0.000107 0.271606 1.280166
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9888 0.1732 5.71 4.2e-08 ***
## LastAuthorFemale1 -0.0393 0.0864 -0.46 0.649
## Year1997 0.4324 0.2299 1.88 0.061 .
## Year1998 0.4107 0.2033 2.02 0.045 *
## Year1999 0.1011 0.2226 0.45 0.650
## Year2000 0.2125 0.2454 0.87 0.388
## Year2001 0.1146 0.2528 0.45 0.651
## Year2002 0.2666 0.2184 1.22 0.224
## Year2003 0.0908 0.2446 0.37 0.711
## Year2004 0.1043 0.2409 0.43 0.666
## Year2005 0.3526 0.1827 1.93 0.055 .
## Year2006 0.2240 0.1966 1.14 0.256
```

```

## Year2007          0.2591      0.1967      1.32      0.189
## Year2008          0.1060      0.2084      0.51      0.611
## Year2009          0.1543      0.1853      0.83      0.406
## Year2010          0.1809      0.1816      1.00      0.321
## Year2011          0.2974      0.1939      1.53      0.127
## Year2012          0.1230      0.1983      0.62      0.536
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.391
## Multiple R-squared:  0.0671, Adjusted R-squared:  -0.0138
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 185 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.262  0.868  0.940  0.889  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      4.67e-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: 214"
## [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
##   28   33   28   31   38   29   22   28   23   30   33   36   34   33   40
## 2011 2012
##   35   38
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   16   12    9   18   19    4   15   17    8   13   20   20   21   22   29
## 2011 2012

```

```
## 20 20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 15 12 8 16 15 4 9 12 7 12 14 15 19 20 22
## 2011 2012
## 17 18
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.5
```



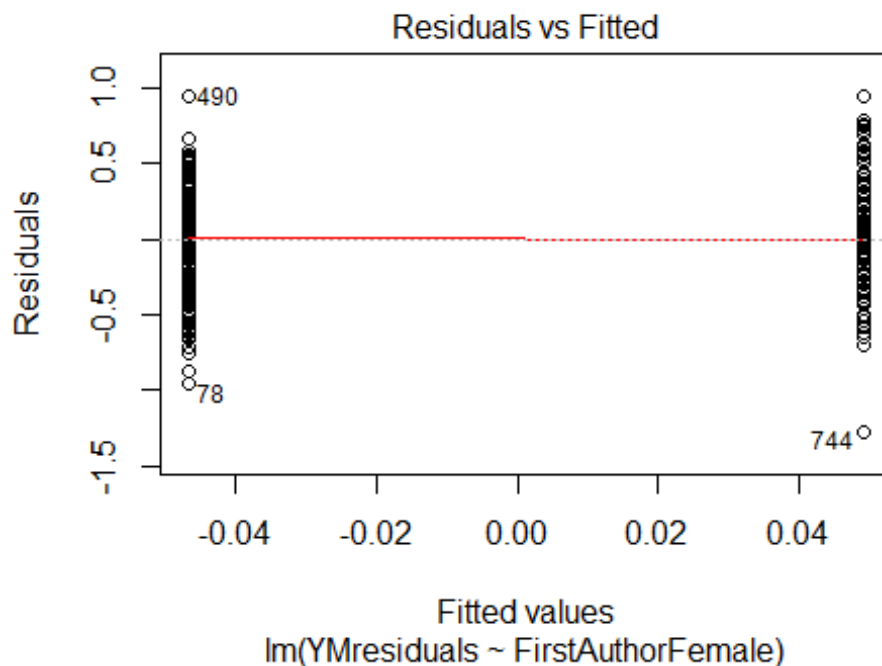
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.39, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 6.64192479699052"
## [1] "Male first author team size 2018 geometric mean: 5.20128756078266"

## 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 = 22, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.89897948556636"
## [1] "Male last author team size 2018 geometric mean: 6.14341916021415"

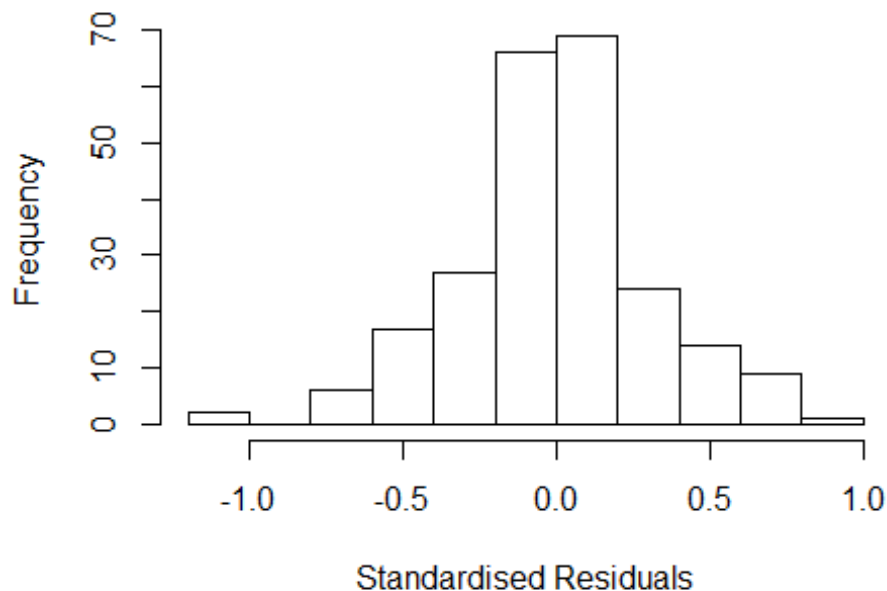
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8.5, 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	2.297	1	1.515
LastAuthorFemale	1.470	1	1.213
UniqueAuthors	14.178	4	1.393
Year	22.543	16	1.102

Residuals from first and last author and team size



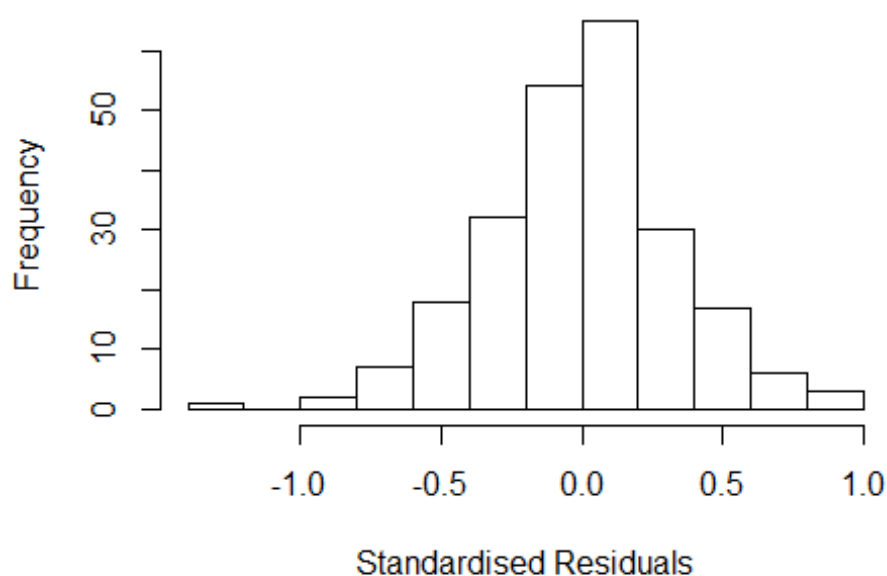
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.18845 -0.15772 -0.00688 0.16555 0.80184
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09651 0.10768 10.18 <2e-16 ***
## FirstAuthorFemale1 0.05715 0.05034 1.14 0.2576
## LastAuthorFemale1 0.00973 0.05434 0.18 0.8581
## UniqueAuthors2 0.23375 0.09487 2.46 0.0145 *
## UniqueAuthors3 0.01586 0.10378 0.15 0.8786
## UniqueAuthors4 0.04398 0.10114 0.43 0.6641
## UniqueAuthors5 0.24130 0.08968 2.69 0.0077 **
## Year1997 0.11245 0.12064 0.93 0.3523
## Year1998 -0.13601 0.11194 -1.22 0.2257
## Year1999 -0.14870 0.11183 -1.33 0.1851
```

```

## Year2000      -0.06843    0.11062   -0.62    0.5369
## Year2001      -0.52454    0.21393   -2.45    0.0150 *
## Year2002      -0.33835    0.12941   -2.61    0.0096 **
## Year2003       0.09873    0.11929    0.83    0.4088
## Year2004      -0.32577    0.13988   -2.33    0.0208 *
## Year2005      -0.22092    0.11476   -1.93    0.0556 .
## Year2006      -0.34306    0.12347   -2.78    0.0059 **
## Year2007      -0.18968    0.13136   -1.44    0.1502
## Year2008      -0.33195    0.10193   -3.26    0.0013 **
## Year2009      -0.23590    0.11440   -2.06    0.0404 *
## Year2010      -0.20393    0.12597   -1.62    0.1070
## Year2011      -0.08213    0.09529   -0.86    0.3897
## Year2012      -0.06360    0.12779   -0.50    0.6192
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.277
## Multiple R-squared:  0.245, Adjusted R-squared:  0.167
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 224 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0254 0.8460 0.9660 0.8870 0.9900 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          4.26e-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 1.738 1          1.318
## LastAuthorFemale  1.462 1          1.209
## Year              2.457 16          1.028

```


Residuals from first and last author



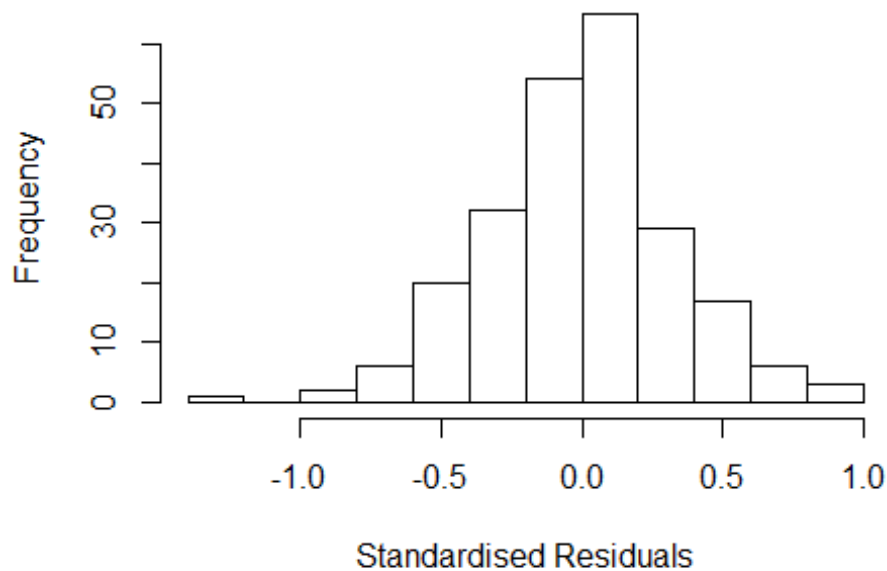
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2948 -0.2091  0.0144  0.1897  0.9298
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21459    0.08356   14.54  <2e-16 ***
## FirstAuthorFemale1  0.06556    0.04995    1.31   0.191
## LastAuthorFemale1  0.02319    0.05652    0.41   0.682
## Year1997         0.02357    0.12128    0.19   0.846
## Year1998        -0.21123    0.11033   -1.91   0.057 .
## Year1999        -0.16556    0.10747   -1.54   0.125
## Year2000        -0.00894    0.10923   -0.08   0.935
## Year2001        -0.48447    0.28622   -1.69   0.092 .
## Year2002        -0.27494    0.13936   -1.97   0.050 *
## Year2003         0.04586    0.11759    0.39   0.697
## Year2004        -0.31151    0.17158   -1.82   0.071 .
## Year2005        -0.18684    0.11526   -1.62   0.106
```

```

## Year2006      -0.29451    0.12216   -2.41    0.017 *
## Year2007      -0.18215    0.13151   -1.39    0.167
## Year2008      -0.28937    0.11537   -2.51    0.013 *
## Year2009      -0.20332    0.11747   -1.73    0.085 .
## Year2010      -0.15479    0.13246   -1.17    0.244
## Year2011      -0.04367    0.10443   -0.42    0.676
## Year2012       0.01463    0.11513    0.13    0.899
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.302
## Multiple R-squared:  0.144, Adjusted R-squared:  0.0726
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 213 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0272 0.8450 0.9500 0.8900 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      4.26e-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 1.684 1      1.298
## Year              1.684 16      1.016

```

Residuals from first author



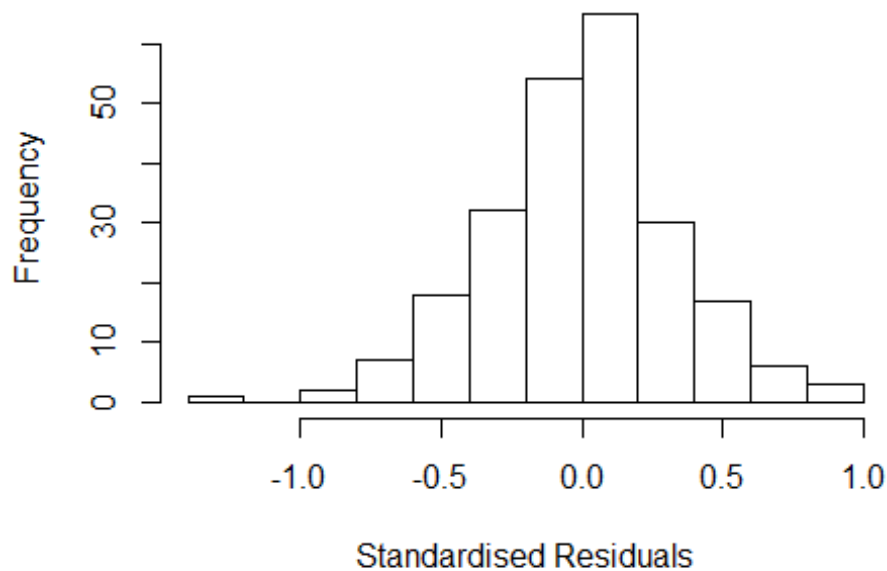
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3061 -0.2092  0.0109  0.1889  0.9280
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.21654    0.08430   14.43  <2e-16 ***
## FirstAuthorFemale1 0.06767    0.04930    1.37   0.171
## Year1997         0.02342    0.12288    0.19   0.849
## Year1998        -0.21055    0.11067   -1.90   0.058 .
## Year1999        -0.16360    0.10796   -1.52   0.131
## Year2000        -0.00838    0.11025   -0.08   0.940
## Year2001        -0.48497    0.29156   -1.66   0.098 .
## Year2002        -0.27702    0.13840   -2.00   0.047 *
## Year2003         0.04621    0.11813    0.39   0.696
## Year2004        -0.30633    0.16883   -1.81   0.071 .
## Year2005        -0.18583    0.11477   -1.62   0.107
## Year2006        -0.29284    0.12296   -2.38   0.018 *
```

```

## Year2007      -0.18295    0.13186   -1.39    0.167
## Year2008      -0.28404    0.11332   -2.51    0.013 *
## Year2009      -0.19873    0.11660   -1.70    0.090 .
## Year2010      -0.14755    0.13198   -1.12    0.265
## Year2011      -0.03539    0.10239   -0.35    0.730
## Year2012       0.02193    0.11567    0.19    0.850
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.3
## Multiple R-squared:  0.143, Adjusted R-squared:  0.0763
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 213 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0191 0.8430 0.9490 0.8890 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      4.26e-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 1.455 1      1.206
## Year      1.455 16      1.012

```

Residuals from last author



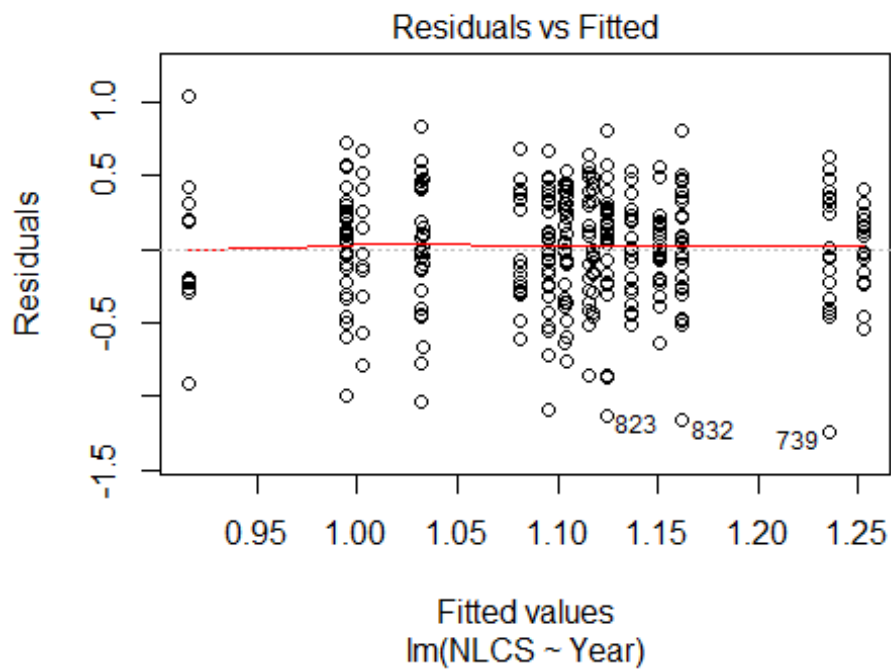
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2608 -0.1886 0.0039 0.1928 0.8887
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.23214 0.07917 15.56 <2e-16 ***
## LastAuthorFemale1 0.03140 0.05527 0.57 0.571
## Year1997 0.05810 0.11687 0.50 0.620
## Year1998 -0.20218 0.10900 -1.85 0.065 .
## Year1999 -0.15592 0.10594 -1.47 0.143
## Year2000 -0.00774 0.10564 -0.07 0.942
## Year2001 -0.50066 0.29001 -1.73 0.086 .
## Year2002 -0.27213 0.13477 -2.02 0.045 *
## Year2003 0.03814 0.11400 0.33 0.738
## Year2004 -0.31376 0.17181 -1.83 0.069 .
## Year2005 -0.17240 0.11248 -1.53 0.127
## Year2006 -0.27302 0.11673 -2.34 0.020 *
```

```

## Year2007          -0.14515      0.12330    -1.18      0.240
## Year2008          -0.27185      0.11447    -2.37      0.018 *
## Year2009          -0.18804      0.11718    -1.60      0.110
## Year2010          -0.14289      0.13093    -1.09      0.276
## Year2011          -0.02477      0.09972    -0.25      0.804
## Year2012           0.02865      0.11633      0.25      0.806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.3
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0696
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 206 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0391 0.8410 0.9520 0.8840 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      4.26e-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: 235"
## [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
##   43   37   26   36   48   45   32   37   26   43   48   51   36   35   32
## 2011 2012
##   38   41
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   18   20    8   13   17   11   11   20   21   25   31   33   23   23   22
## 2011 2012

```

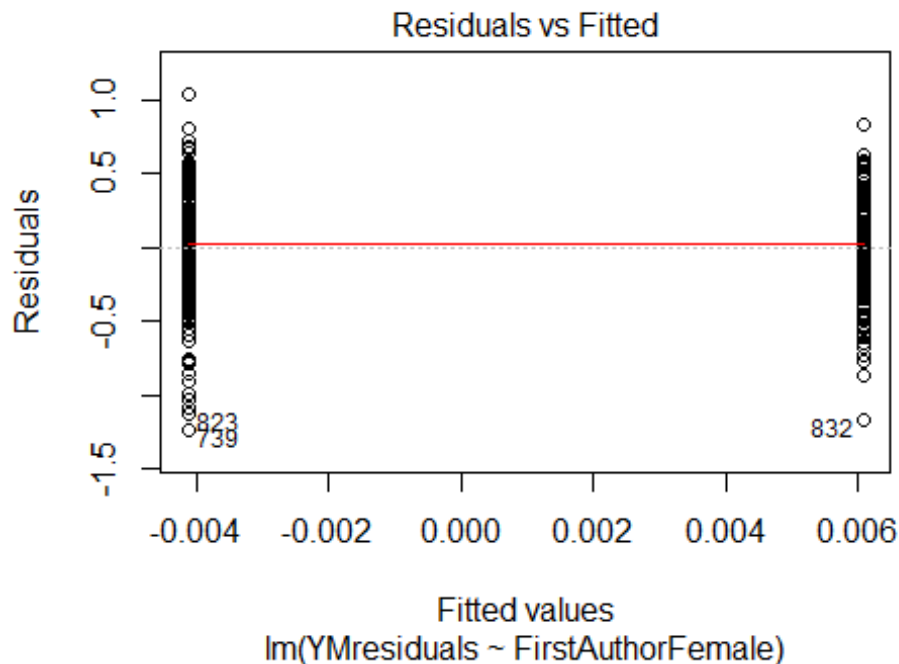
```
## 30 28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 16 6 12 14 9 9 16 18 19 26 28 23 21 20
## 2011 2012
## 29 28
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 26, df = 16, p-value = 0.05
```



```
##
## 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: 4.09324589064645"
## [1] "Male first author team size 2018 geometric mean: 3.53230485861734"
## 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 = 45, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.61789292741588"
## [1] "Male last author team size 2018 geometric mean: 3.52891281418207"

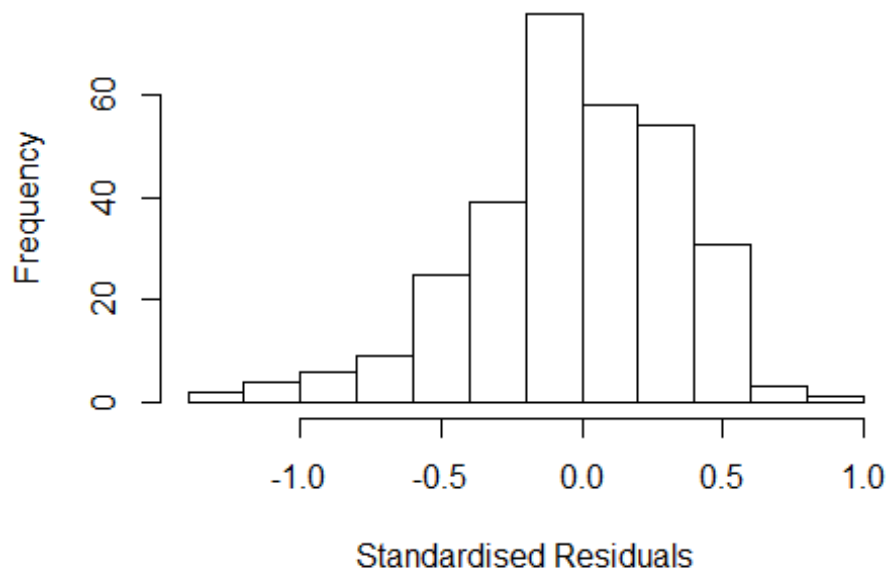
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 42, 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	1.507	1	1.228
LastAuthorFemale	1.266	1	1.125
UniqueAuthors	3.909	4	1.186
Year	6.013	16	1.058

Residuals from first and last author and team size



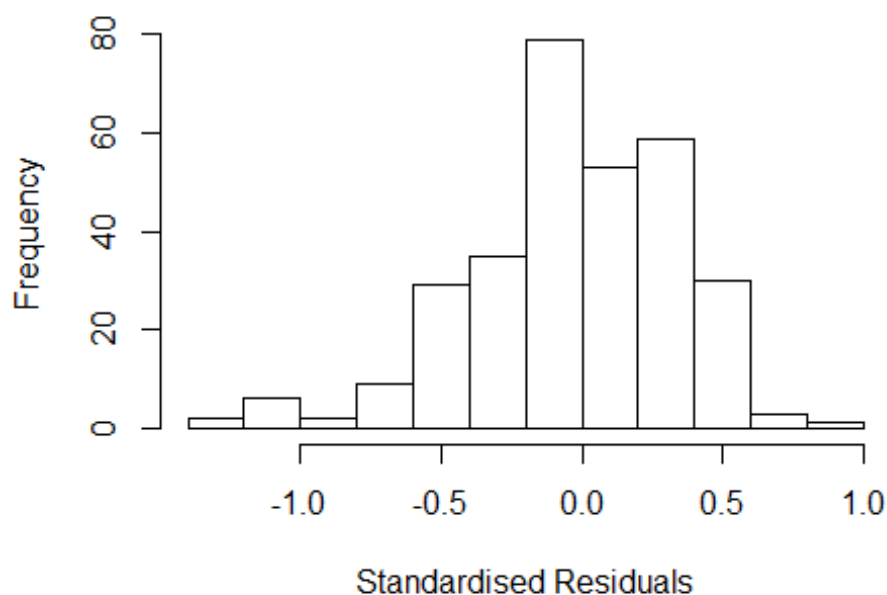
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2753 -0.2275 -0.0139 0.2365 0.8252
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.131583 0.166946 6.78 7e-11 ***
## FirstAuthorFemale1 0.000636 0.047053 0.01 0.989
## LastAuthorFemale1 0.073357 0.060171 1.22 0.224
## UniqueAuthors2 0.114065 0.116991 0.97 0.330
## UniqueAuthors3 0.057584 0.114727 0.50 0.616
## UniqueAuthors4 0.096769 0.110491 0.88 0.382
## UniqueAuthors5 0.107131 0.114706 0.93 0.351
## Year1997 -0.091759 0.149081 -0.62 0.539
## Year1998 -0.200924 0.203366 -0.99 0.324
## Year1999 -0.075446 0.160435 -0.47 0.639
```

```

## Year2000      -0.102938    0.172071   -0.60    0.550
## Year2001      -0.445989    0.180522   -2.47    0.014 *
## Year2002      -0.306363    0.221088   -1.39    0.167
## Year2003       0.017852    0.143812    0.12    0.901
## Year2004      -0.092257    0.187305   -0.49    0.623
## Year2005      -0.086524    0.154111   -0.56    0.575
## Year2006      -0.120051    0.155371   -0.77    0.440
## Year2007      -0.158958    0.146490   -1.09    0.279
## Year2008      -0.194381    0.165755   -1.17    0.242
## Year2009      -0.091913    0.147256   -0.62    0.533
## Year2010       0.046923    0.159312    0.29    0.769
## Year2011      -0.028828    0.149130   -0.19    0.847
## Year2012      -0.058194    0.150891   -0.39    0.700
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.0839, Adjusted R-squared:  0.0132
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 272 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.144  0.868  0.944  0.893  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      3.25e-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 1.416 1      1.190
## LastAuthorFemale  1.227 1      1.108
## Year              1.632 16      1.015

```

Residuals from first and last author



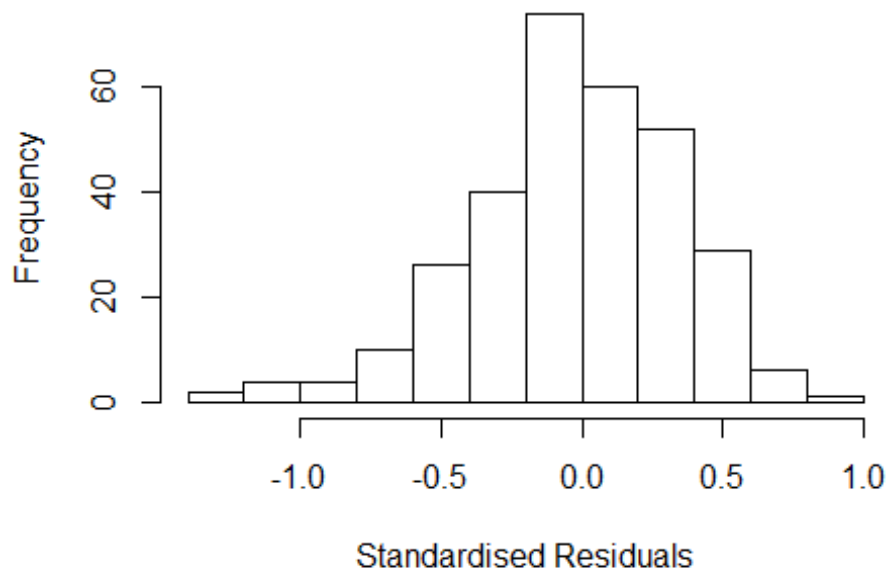
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2527 -0.2191 -0.0188  0.2406  0.8097
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.20882    0.14111   8.57 6.5e-16 ***
## FirstAuthorFemale1  0.00659    0.04641   0.14  0.887
## LastAuthorFemale1  0.07461    0.06044   1.23  0.218
## Year1997         -0.08774    0.15375  -0.57  0.569
## Year1998         -0.17546    0.19649  -0.89  0.373
## Year1999         -0.07145    0.16547  -0.43  0.666
## Year2000         -0.10293    0.18105  -0.57  0.570
## Year2001         -0.44622    0.18292  -2.44  0.015 *
## Year2002         -0.32006    0.21561  -1.48  0.139
## Year2003          0.03456    0.14951   0.23  0.817
## Year2004         -0.10015    0.18896  -0.53  0.597
## Year2005         -0.07784    0.15861  -0.49  0.624
```

```

## Year2006      -0.10190    0.15721   -0.65    0.517
## Year2007      -0.14161    0.15130   -0.94    0.350
## Year2008      -0.18264    0.16966   -1.08    0.283
## Year2009      -0.08257    0.15149   -0.55    0.586
## Year2010       0.04390    0.16391    0.27    0.789
## Year2011      -0.01364    0.15220   -0.09    0.929
## Year2012      -0.05353    0.15738   -0.34    0.734
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.0768, Adjusted R-squared:  0.0194
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 270 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.161  0.854  0.944  0.891  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      3.25e-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 1.431 1      1.196
## Year      1.431 16      1.011

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2749 -0.2265 -0.0157 0.2404 0.8109
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2242 0.1434 8.54 7.9e-16 ***
## FirstAuthorFemale1 0.0108 0.0462 0.23 0.816
## Year1997 -0.1042 0.1555 -0.67 0.503
## Year1998 -0.1756 0.1997 -0.88 0.380
## Year1999 -0.0774 0.1710 -0.45 0.651
## Year2000 -0.1185 0.1839 -0.64 0.520
## Year2001 -0.4436 0.1863 -2.38 0.018 *
## Year2002 -0.3193 0.2152 -1.48 0.139
## Year2003 0.0233 0.1514 0.15 0.878
## Year2004 -0.0985 0.1901 -0.52 0.605
## Year2005 -0.0881 0.1595 -0.55 0.581
## Year2006 -0.0945 0.1584 -0.60 0.551
```

```

## Year2007          -0.1511      0.1539   -0.98    0.327
## Year2008          -0.1778      0.1742   -1.02    0.308
## Year2009          -0.0906      0.1535   -0.59    0.556
## Year2010           0.0507      0.1652    0.31    0.759
## Year2011          -0.0139      0.1533   -0.09    0.928
## Year2012          -0.0417      0.1604   -0.26    0.795
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.0699, Adjusted R-squared:  0.0153
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 272 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.146  0.860  0.943  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      3.25e-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 1.199 1      1.095
## Year              1.199 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
## -1.2575 -0.2205 -0.0186  0.2396  0.8059

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2120    0.1407    8.61 4.7e-16 ***
## LastAuthorFemale1  0.0751    0.0601    1.25  0.213
## Year1997          -0.0892    0.1543   -0.58  0.564
## Year1998          -0.1769    0.1957   -0.90  0.367
## Year1999          -0.0732    0.1658   -0.44  0.659
## Year2000          -0.1030    0.1813   -0.57  0.570
## Year2001          -0.4467    0.1831   -2.44  0.015 *
## Year2002          -0.3208    0.2158   -1.49  0.138
## Year2003           0.0324    0.1494    0.22  0.829
## Year2004          -0.0987    0.1877   -0.53  0.600
## Year2005          -0.0772    0.1588   -0.49  0.627
## Year2006          -0.1022    0.1574   -0.65  0.517
## Year2007          -0.1418    0.1513   -0.94  0.350
## Year2008          -0.1818    0.1688   -1.08  0.283
## Year2009          -0.0832    0.1519   -0.55  0.584
## Year2010           0.0455    0.1633    0.28  0.781
## Year2011          -0.0133    0.1522   -0.09  0.931
## Year2012          -0.0529    0.1573   -0.34  0.737
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.344
## Multiple R-squared:  0.0771, Adjusted R-squared:  0.023
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 270 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.153  0.854  0.945  0.890  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      3.25e-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: 308"
## [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
##    23   13   17   13   13    6   16   13    7   17   12   10    9   14   13
## 2011 2012
##    12   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     8    3    4    4    5    1    4    4    3    5    3    5    3    4    4
## 2011 2012
##     6    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     5    2    3    4    4    1    3    3    3    5    3    4    2    2    4
## 2011 2012
##     6    7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.91486764116886"
## [1] "Male first author team size 2018 geometric mean: 3.1291346445319"

## 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 = 19, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male last author team size 2018 geometric mean: 3.55198718249027"

## 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 = 7.5, 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"

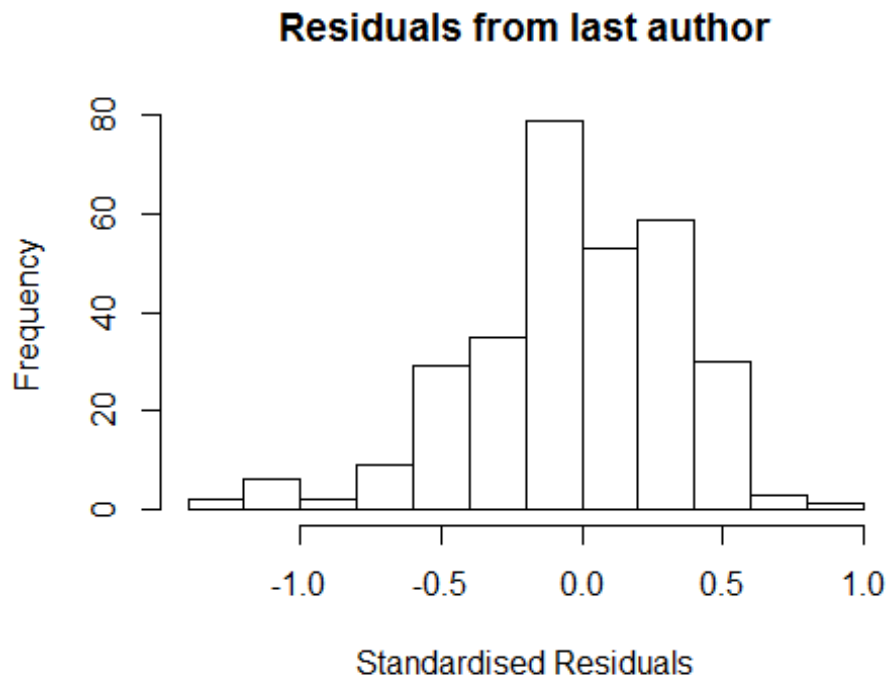
```



```
## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate

## [1] "Regression 2: First author gender, Last 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 3: First author gender, Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 16.81  1         4.100
## Year              16.81 16         1.092

## [1] "List of 0 outliers 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
## -0.85797 -0.15173 -0.00391  0.17539  0.90621
##
```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4689    0.1043   14.09 < 2e-16 ***
## FirstAuthorFemale1 0.0342    0.1140    0.30 0.76531
## Year1997          0.0090    0.1218    0.07 0.94144
## Year1998         -0.1378    0.6318   -0.22 0.82841
## Year1999         -0.6452    0.3309   -1.95 0.05775 .
## Year2000         -0.1000    0.1694   -0.59 0.55830
## Year2001          0.1261    0.1043    1.21 0.23306
## Year2002         -0.5495    0.2338   -2.35 0.02341 *
## Year2003         -0.0240    0.1141   -0.21 0.83457
## Year2004         -0.1395    0.1360   -1.03 0.31064
## Year2005         -0.3973    0.1621   -2.45 0.01839 *
## Year2006         -0.4953    0.3760   -1.32 0.19470
## Year2007         -0.6107    0.1645   -3.71 0.00059 ***
## Year2008         -0.4064    0.1783   -2.28 0.02769 *
## Year2009         -0.6050    0.2307   -2.62 0.01202 *
## Year2010         -0.2133    0.2430   -0.88 0.38486
## Year2011         -0.4878    0.1539   -3.17 0.00281 **
## Year2012         -0.4603    0.1863   -2.47 0.01749 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.349, Adjusted R-squared:  0.0916
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.500  0.908  0.968  0.922  0.994  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00      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-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 lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate

## [1] "Sample size for the above analysis: 61"
## [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 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    3    4    8    4    3    2    3    3    7    9    4    7    11    12
## 2012
##    6
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    1    0    2    0    2    1    3    3    5    6    2    2    7    6
## 2012
##    6
##
## 1996 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    1    0    1    0    2    1    3    3    5    6    2    2    7    5
## 2012
##    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.57303604546244"
## [1] "Male first author team size 2018 geometric mean: 3"

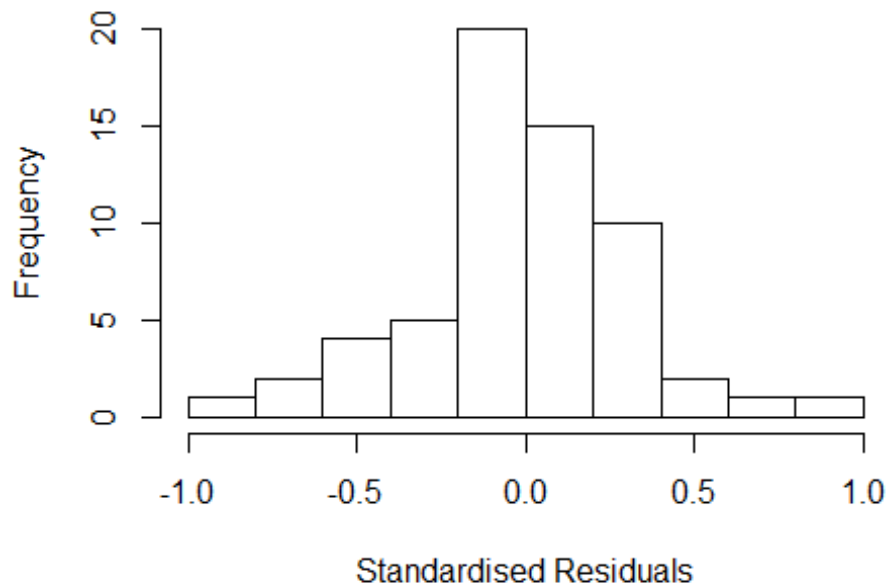
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.50630624679114"
## [1] "Male last author team size 2018 geometric mean: 3"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

Residuals from first author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8, 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 6.495e+02 1          25.48
## LastAuthorFemale -1.019e+02 1           NaN
## UniqueAuthors    -4.042e+45 4           NaN
## Year              -1.595e+48 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 +
UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

##           Min           1Q           Median           3Q           Max
## -7.12e-01 -8.45e-02 -1.28e-15  8.47e-02  4.54e-01
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4899     0.1439   10.35 1.6e-10 ***
## FirstAuthorFemale1 -0.0353     0.0992   -0.36 0.72470
## LastAuthorFemale1  0.1182     0.1129    1.05 0.30515
## UniqueAuthors2     -0.8499     0.1976   -4.30 0.00023 ***
## UniqueAuthors3     -0.6529     0.1466   -4.45 0.00015 ***
## UniqueAuthors4     -0.6319     0.1439   -4.39 0.00018 ***
## UniqueAuthors5     -0.6222     0.1393   -4.47 0.00015 ***
## Year2000           0.5933     0.1399    4.24 0.00027 ***
## Year2002           0.4379     0.1546    2.83 0.00898 **
## Year2003           0.5440     0.0943    5.77 5.1e-06 ***
## Year2004           0.2099     0.6743    0.31 0.75814
## Year2005           0.2977     0.1766    1.69 0.10432
## Year2006           0.1845     0.1436    1.28 0.21084
## Year2007           0.0469     0.1645    0.28 0.77813
## Year2008           0.3191     0.0796    4.01 0.00048 ***
## Year2009           0.1653     0.0994    1.66 0.10888
## Year2010           0.0361     0.1439    0.25 0.80386
## Year2011           0.3231     0.0909    3.56 0.00153 **
## Year2012           0.0559     0.1633    0.34 0.73497
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.258
## Multiple R-squared:  0.401, Adjusted R-squared:  -0.0309
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 28 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.429  0.827  0.936  0.882  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      2.27e-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"

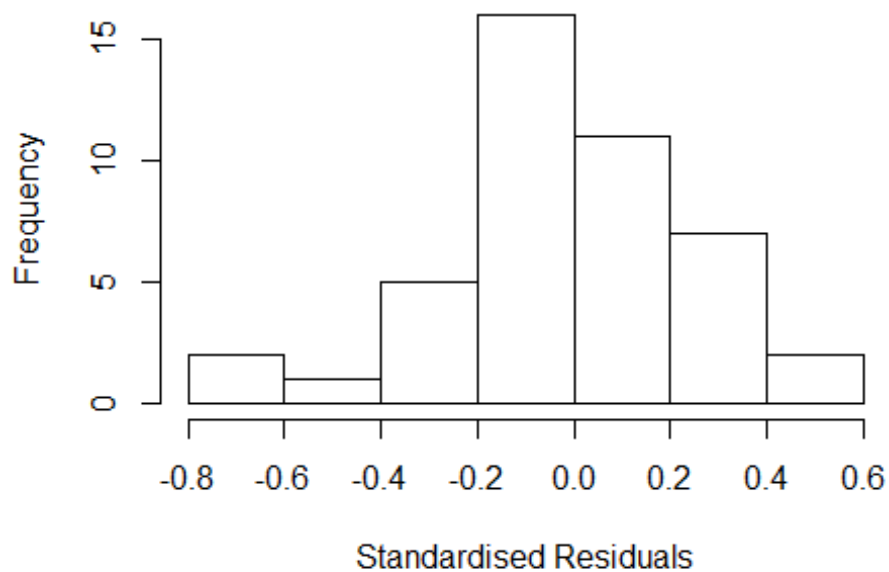
```

```
## 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

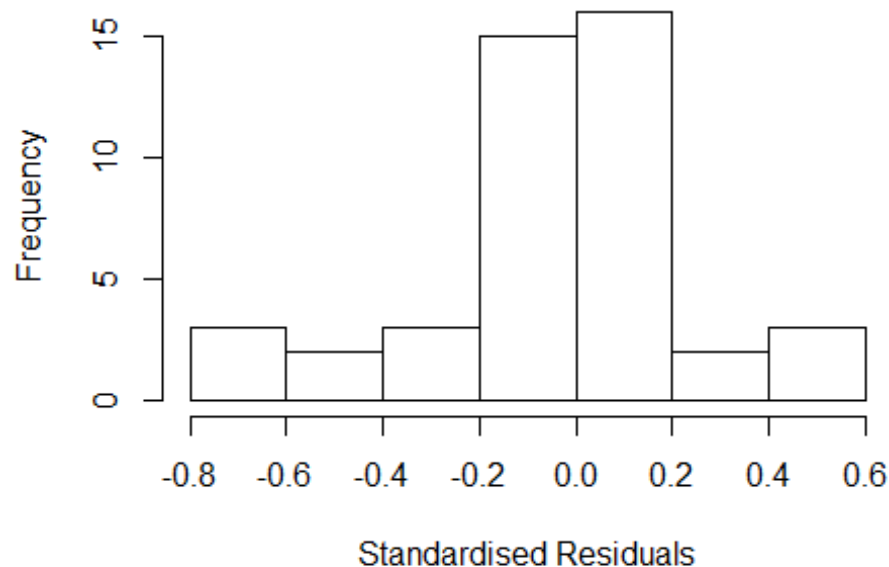
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author and team size



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1          NaN
## LastAuthorFemale  NaN 1          NaN
## Year              NaN 12         NaN
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -6.88e-01 -7.36e-02  8.33e-16  9.54e-02  5.67e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    8.58e-01   0.00e+00      Inf < 2e-16 ***
## FirstAuthorFemale1 -9.89e-02   8.81e-02  -1.12e+00   0.271
## LastAuthorFemale1  7.89e-02   1.05e-01   7.50e-01   0.460
## Year2000         6.03e-01   6.83e-09  8.82e+07 < 2e-16 ***
## Year2002         4.79e-01   6.82e-02  7.03e+00  9.9e-08 ***
## Year2003         5.23e-01   0.00e+00      Inf < 2e-16 ***
## Year2004         2.99e-01   5.22e-01   5.70e-01   0.571
## Year2005         3.45e-01   1.49e-01   2.32e+00   0.028 *
## Year2006         1.94e-01   1.51e-01   1.29e+00   0.208
## Year2007         9.29e-02   1.24e-01   7.50e-01   0.461
## Year2008         2.61e-01   1.09e-01   2.39e+00   0.023 *
## Year2009         2.18e-01   8.93e-02   2.45e+00   0.021 *
```

```

## Year2010          1.01e-01  1.54e-01  6.50e-01  0.519
## Year2011          3.10e-01  5.54e-02  5.60e+00  4.8e-06 ***
## Year2012          9.67e-02  1.84e-01  5.20e-01  0.604
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.257
## Multiple R-squared:  0.266, Adjusted R-squared:  -0.0877
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 35 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.454  0.814  0.980  0.886  0.993  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      2.27e-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 38.3  1          6.189
## Year              38.3 12          1.164
##
## [1] "List of 0 outliers 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
## -0.6829 -0.0605  0.0000  0.0885  0.5300
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)

```



```

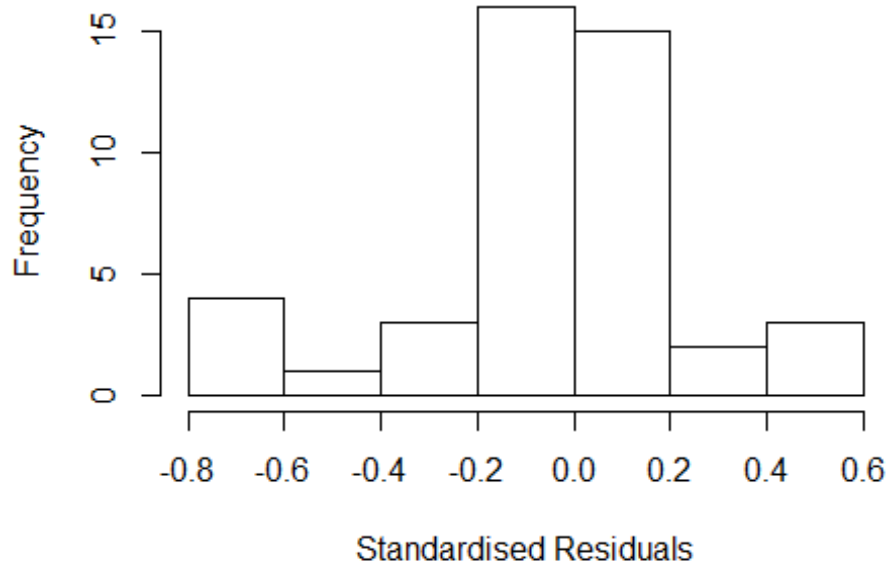
## (Intercept)      8.58e-01  1.59e-08  5.40e+07  < 2e-16 ***
## FirstAuthorFemale1 -1.06e-01  8.62e-02 -1.23e+00  0.230
## Year2000          6.03e-01  7.01e-09  8.60e+07  < 2e-16 ***
## Year2002          4.83e-01  6.57e-02  7.34e+00  3.5e-08 ***
## Year2003          5.23e-01  1.48e-08  3.54e+07  < 2e-16 ***
## Year2004          3.80e-01  4.18e-01  9.10e-01  0.371
## Year2005          3.80e-01  1.45e-01  2.62e+00  0.014 *
## Year2006          2.48e-01  1.13e-01  2.20e+00  0.036 *
## Year2007          1.14e-01  1.16e-01  9.80e-01  0.334
## Year2008          3.04e-01  1.18e-01  2.58e+00  0.015 *
## Year2009          2.25e-01  8.74e-02  2.58e+00  0.015 *
## Year2010          1.38e-01  1.22e-01  1.13e+00  0.268
## Year2011          3.28e-01  5.86e-02  5.60e+00  4.3e-06 ***
## Year2012          1.17e-01  2.10e-01  5.60e-01  0.582
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.242
## Multiple R-squared:  0.266, Adjusted R-squared:  -0.0517
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.405  0.806  0.968  0.866  0.993  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

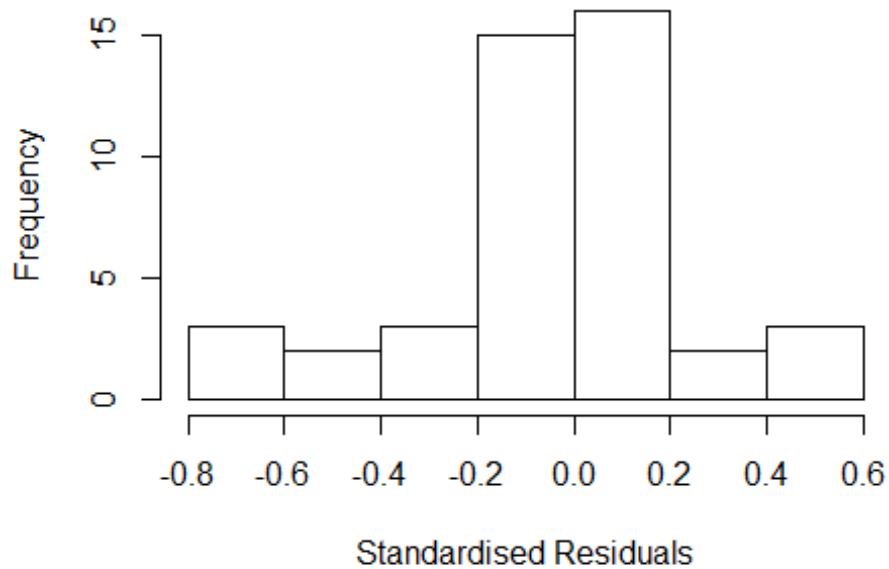
```

Residuals from first author



##	GVIF	Df	$GVIF^{1/(2 \cdot Df)}$
## LastAuthorFemale	NaN	1	NaN
## Year	NaN	12	NaN

Residuals from last author

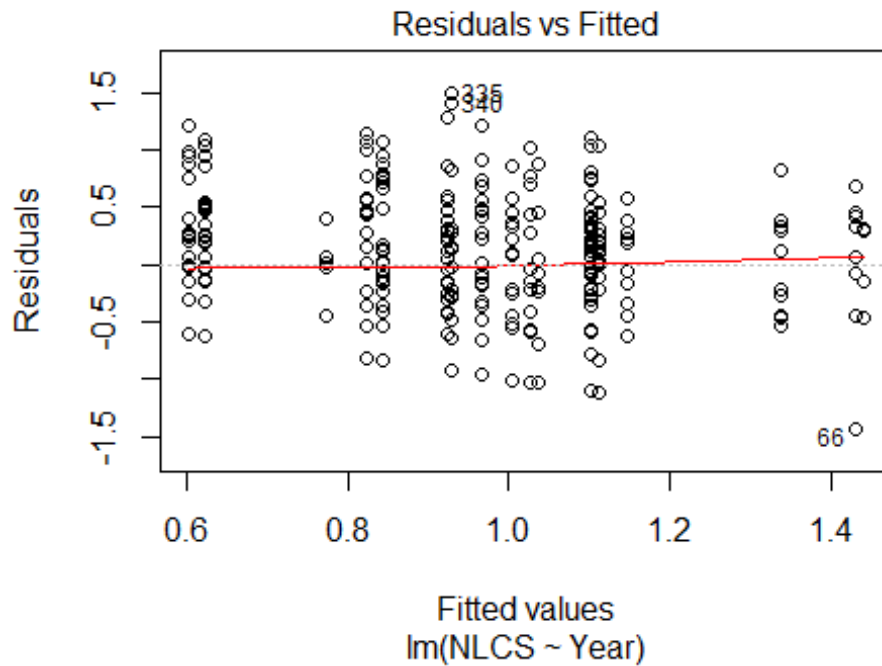


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.6707 -0.1055 0.0156 0.1211 0.6093
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8580 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 0.0912 0.1074 0.85 0.402
## Year2000 0.6030 0.0000 Inf < 2e-16 ***
## Year2002 0.4300 0.0893 4.82 3.9e-05 ***
## Year2003 0.5230 0.0000 Inf < 2e-16 ***
## Year2004 0.1581 0.4457 0.35 0.725
## Year2005 0.2743 0.1597 1.72 0.096 .
## Year2006 0.1300 0.1512 0.86 0.397
## Year2007 0.0429 0.1008 0.43 0.674
## Year2008 0.2059 0.1255 1.64 0.111
## Year2009 0.1195 0.0145 8.24 3.4e-09 ***
## Year2010 0.0587 0.1814 0.32 0.749
## Year2011 0.2680 0.0381 7.03 8.2e-08 ***
## Year2012 0.0320 0.1450 0.22 0.827
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.284
## Multiple R-squared: 0.232, Adjusted R-squared: -0.1
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 33 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.557 0.866 0.970 0.896 0.990 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 2.27e-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: 44"
## [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
##    20    30    29    35    33    43    33    31    50    31    68    40    72    59    49
## 2011 2012
##    52    67
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     5     8     9    15    10     4    10    17    18    14    31    25    37    34    33
## 2011 2012
##    28    43
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     5     6     8    13     9     4     6    11    14    13    23    21    28    28    29
## 2011 2012
##    24    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 = 12, df = 16, p-value = 0.7

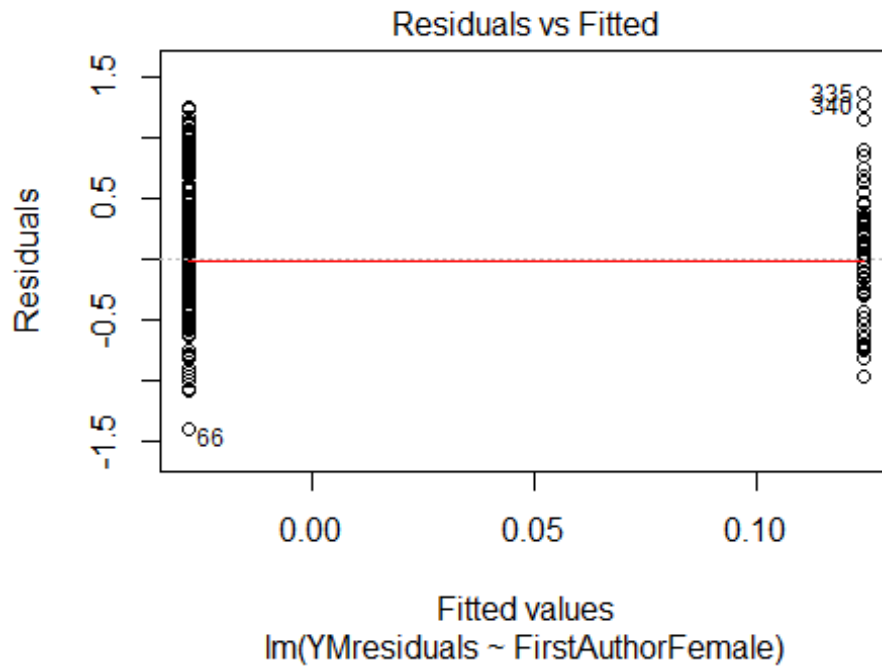
```



```
##
## 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: 3.33702088205365"
## [1] "Male first author team size 2018 geometric mean: 2.89950533163095"

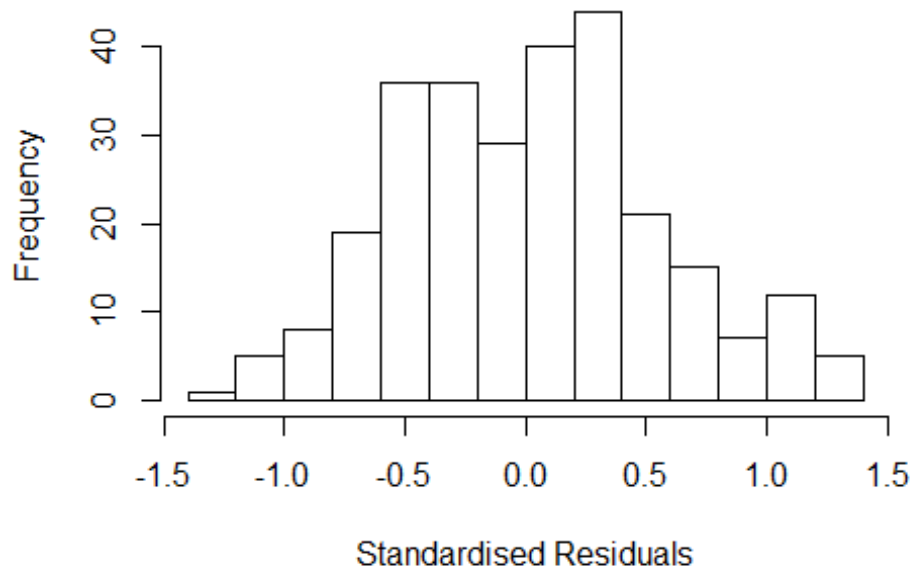
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 90, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 2.99616284871195"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF^(1/(2*Df))
FirstAuthorFemale	1.494	1	1.222
LastAuthorFemale	1.349	1	1.162
UniqueAuthors	3.302	4	1.161
Year	4.818	16	1.050

Residuals from first and last author and team size



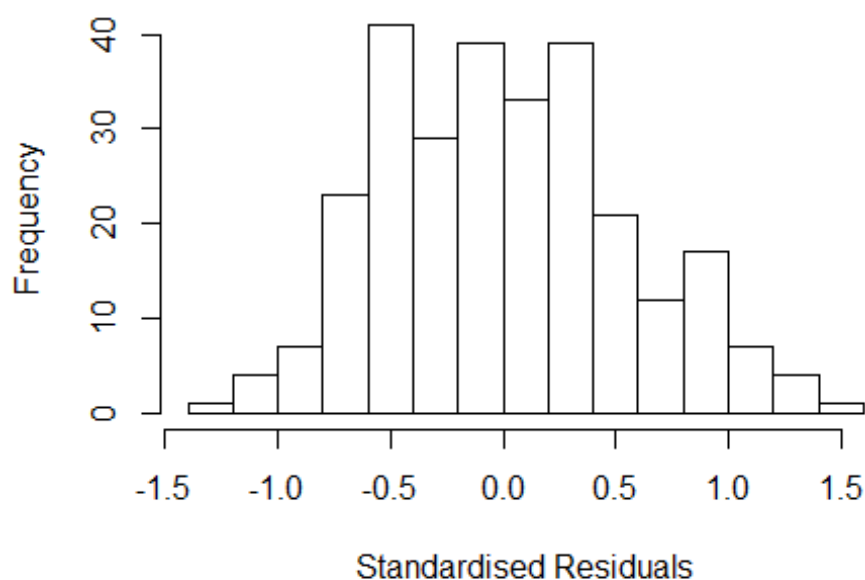
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2686 -0.3968 0.0356 0.3484 1.3577
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7368 0.1553 4.74 3.5e-06 ***
## FirstAuthorFemale1 0.1492 0.0980 1.52 0.1291
## LastAuthorFemale1 0.0601 0.1188 0.51 0.6135
## UniqueAuthors2 0.0247 0.1277 0.19 0.8471
## UniqueAuthors3 -0.1693 0.1246 -1.36 0.1753
## UniqueAuthors4 -0.0142 0.1654 -0.09 0.9317
## UniqueAuthors5 0.0226 0.1887 0.12 0.9047
## Year1997 0.5317 0.3203 1.66 0.0982 .
## Year1998 0.2157 0.2709 0.80 0.4265
## Year1999 0.2534 0.2484 1.02 0.3086
```

```

## Year2000          0.4367      0.1921      2.27      0.0238 *
## Year2001          0.6644      0.2295      2.90      0.0041 **
## Year2002          0.6075      0.2031      2.99      0.0030 **
## Year2003          0.1497      0.3168      0.47      0.6370
## Year2004          0.3078      0.1895      1.62      0.1056
## Year2005          0.2795      0.1932      1.45      0.1491
## Year2006          -0.1578      0.1770     -0.89      0.3734
## Year2007          -0.0167      0.2196     -0.08      0.9394
## Year2008          -0.1806      0.1666     -1.08      0.2795
## Year2009          0.0531      0.1878      0.28      0.7775
## Year2010          0.1883      0.1827      1.03      0.3038
## Year2011          0.1969      0.1785      1.10      0.2709
## Year2012          0.3364      0.1665      2.02      0.0443 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared:  0.159, Adjusted R-squared:  0.0865
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 264 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.522  0.884  0.953   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      3.60e-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 1.373 1      1.172
## LastAuthorFemale  1.378 1      1.174
## Year              1.862 16      1.020

```


Residuals from first and last author



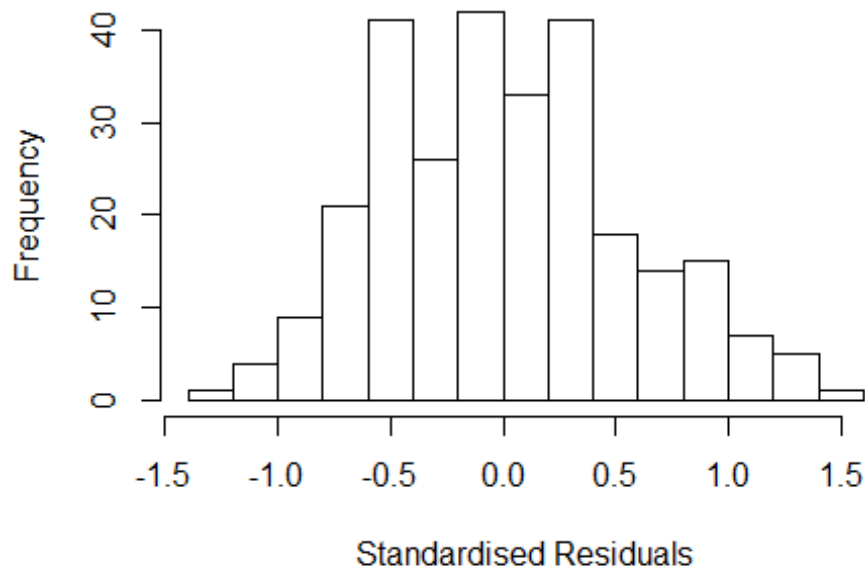
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2951 -0.4182 -0.0164  0.3521  1.4782
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.74425    0.12913   5.76 2.3e-08 ***
## FirstAuthorFemale1 0.14375    0.09569   1.50  0.1343
## LastAuthorFemale1 0.06004    0.12046   0.50  0.6186
## Year1997         0.55088    0.29431   1.87  0.0624 .
## Year1998         0.16186    0.25958   0.62  0.5335
## Year1999         0.19453    0.23807   0.82  0.4146
## Year2000         0.32237    0.18252   1.77  0.0785 .
## Year2001         0.63143    0.19885   3.18  0.0017 **
## Year2002         0.51509    0.20202   2.55  0.0114 *
## Year2003         0.05176    0.29152   0.18  0.8592
## Year2004         0.27943    0.19540   1.43  0.1539
## Year2005         0.19398    0.18944   1.02  0.3068
```

```

## Year2006          -0.22363      0.16373      -1.37      0.1732
## Year2007          -0.05279      0.21790      -0.24      0.8088
## Year2008          -0.23919      0.15661      -1.53      0.1279
## Year2009          -0.00473      0.17505      -0.03      0.9785
## Year2010           0.15394      0.17998       0.86      0.3932
## Year2011           0.15472      0.16760       0.92      0.3568
## Year2012           0.30044      0.15763       1.91      0.0578 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.557
## Multiple R-squared:  0.14,   Adjusted R-squared:  0.0801
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.461  0.865   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      3.60e-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 1.348 1          1.161
## Year              1.348 16          1.009

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3036 -0.4318 -0.0242 0.3475 1.4612
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.74261 0.12942 5.74 2.7e-08 ***
## FirstAuthorFemale1 0.15149 0.09501 1.59 0.1121
## Year1997 0.56101 0.28465 1.97 0.0498 *
## Year1998 0.16274 0.26042 0.62 0.5326
## Year1999 0.20624 0.24009 0.86 0.3911
## Year2000 0.33568 0.18151 1.85 0.0655 .
## Year2001 0.66144 0.19992 3.31 0.0011 **
## Year2002 0.51673 0.20223 2.56 0.0112 *
## Year2003 0.06273 0.29108 0.22 0.8295
## Year2004 0.28684 0.19625 1.46 0.1451
## Year2005 0.20386 0.18842 1.08 0.2803
## Year2006 -0.22292 0.16456 -1.35 0.1767
```

```

## Year2007          -0.04782    0.22438   -0.21    0.8314
## Year2008          -0.22742    0.15338   -1.48    0.1394
## Year2009           0.00507    0.17537    0.03    0.9769
## Year2010           0.16922    0.17721    0.95    0.3405
## Year2011           0.15698    0.16841    0.93    0.3521
## Year2012           0.30924    0.15745    1.96    0.0506 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.556
## Multiple R-squared:  0.139, Adjusted R-squared:  0.0828
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 255 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.470  0.863  0.947  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      3.60e-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 1.282 1      1.132
## Year              1.282 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3717 -0.4321 -0.0241  0.3650  1.6379

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.77484    0.12362   6.27 1.5e-09 ***
## LastAuthorFemale1 0.08392    0.11692   0.72 0.4736
## Year1997        0.59686    0.32224   1.85 0.0651 .
## Year1998        0.14716    0.24580   0.60 0.5499
## Year1999        0.18514    0.22517   0.82 0.4117
## Year2000        0.31293    0.17759   1.76 0.0792 .
## Year2001        0.62546    0.20383   3.07 0.0024 **
## Year2002        0.48450    0.19820   2.44 0.0152 *
## Year2003        0.00524    0.28247   0.02 0.9852
## Year2004        0.27500    0.19655   1.40 0.1630
## Year2005        0.19502    0.19521   1.00 0.3187
## Year2006       -0.22406    0.16326  -1.37 0.1711
## Year2007       -0.04341    0.21661  -0.20 0.8413
## Year2008       -0.23497    0.15503  -1.52 0.1308
## Year2009       -0.00744    0.16993  -0.04 0.9651
## Year2010        0.12977    0.17458   0.74 0.4579
## Year2011        0.15345    0.16435   0.93 0.3513
## Year2012        0.29144    0.15575   1.87 0.0624 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.13, Adjusted R-squared:  0.073
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu. Median   Mean 3rd Qu.   Max.
##   0.387 0.894 0.948 0.913 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      3.60e-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: 278"
## [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
##    4    7    4    9    3    9    7    1    1    3   10    4   17    1    2
## 2011 2012
##    1    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    4    2    3    2    0    1    3    4    2    6    0    1
## 2011 2012
##    0    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    1    1    4    1    3    2    0    0    3    0    2    5    0    1
## 2011 2012
##    0    1
## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 24"
## [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
##    1    8    5    2    5    3    1    1    2    2    5   11    5   15   12
## 2011 2012
##   10   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    5    2    1    2    1    1    0    1    2    2    8    2    9    9
## 2011 2012
##    8   11
##

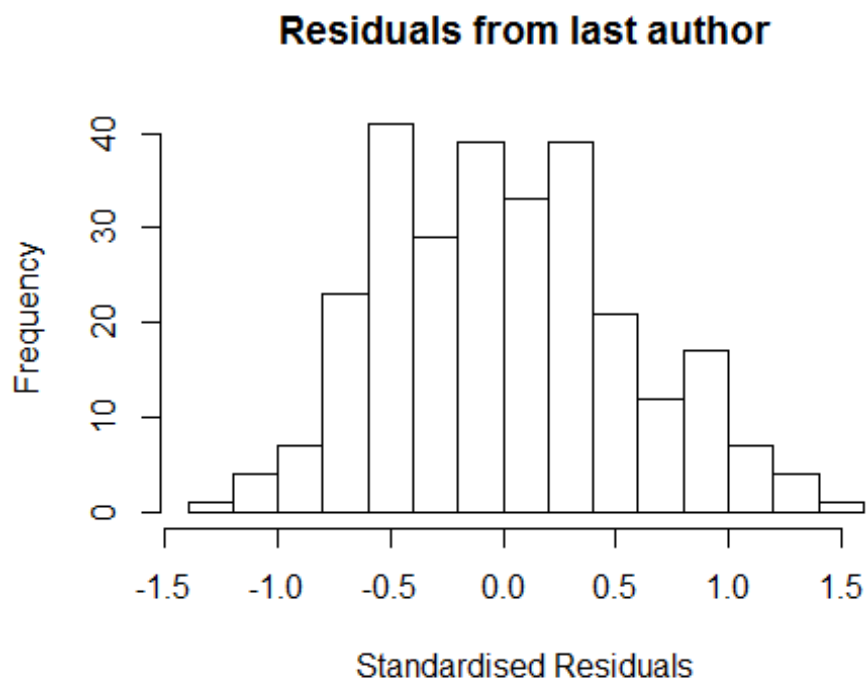
```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    5    2    0    0    1    1    0    1    1    1    7    2    9    9
## 2011 2012
##    7   11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 6.29690351632899"
## [1] "Male first author team size 2018 geometric mean: 5.02377286301916"

## 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 = 20, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 5.69845654698429"

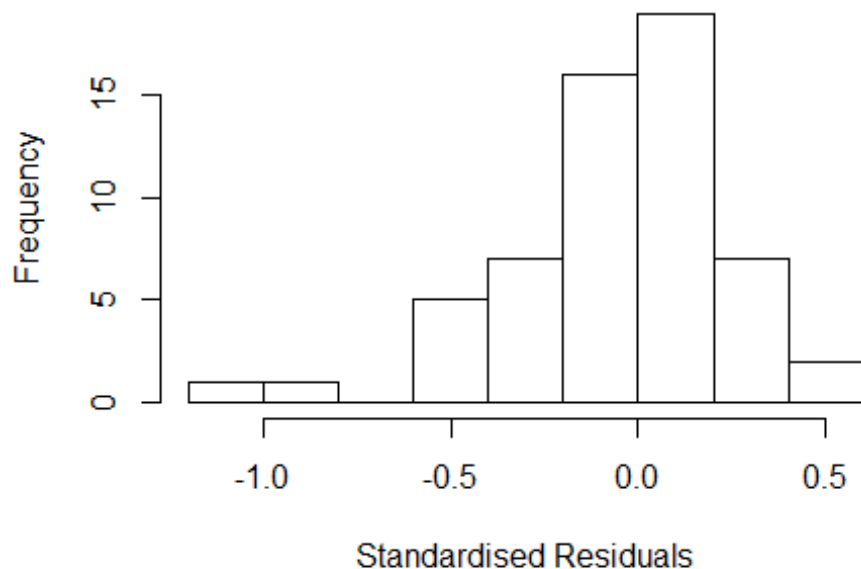
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
```

```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4, p-value = 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 -6.002e+14  1          NaN
## LastAuthorFemale  -1.213e+14  1          NaN
## UniqueAuthors      5.379e+43  4       292643
## Year                1.010e+72 13          588
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.016 -0.175 0.000 0.155 0.459
```



```

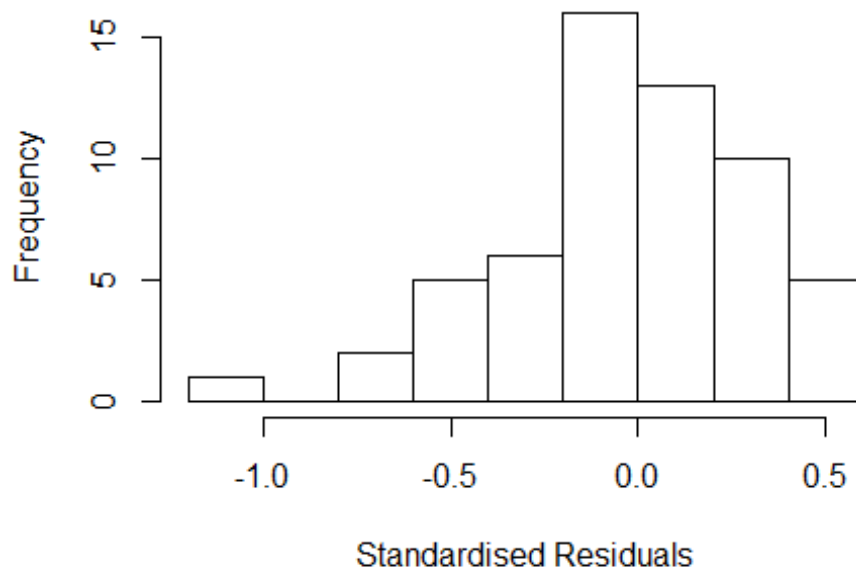
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9127    0.1712    5.33 4.7e-06 ***
## FirstAuthorFemale1 0.2168    0.0887    2.44 0.0193 *
## LastAuthorFemale1 -0.2970    0.0945   -3.14 0.0033 **
## UniqueAuthors2    0.0464    0.1513    0.31 0.7606
## UniqueAuthors3   -0.2395    0.1555   -1.54 0.1319
## UniqueAuthors4   -0.3499    0.1532   -2.28 0.0280 *
## UniqueAuthors5   -0.2140    0.1133   -1.89 0.0665 .
## Year1997          0.0112    0.1725    0.06 0.9487
## Year1998         -0.0469    0.2993   -0.16 0.8763
## Year2001         -0.0302    0.1948   -0.15 0.8778
## Year2002         -0.2912    0.0887   -3.28 0.0022 **
## Year2004         -0.2362    0.1948   -1.21 0.2329
## Year2005          0.6852    0.1954    3.51 0.0012 **
## Year2006          0.3478    0.2327    1.49 0.1433
## Year2007          0.1863    0.1712    1.09 0.2833
## Year2008         -0.0204    0.1661   -0.12 0.9030
## Year2009          0.2970    0.1890    1.57 0.1243
## Year2010          0.4535    0.2043    2.22 0.0325 *
## Year2011          0.3974    0.1687    2.35 0.0238 *
## Year2012          0.4047    0.1771    2.29 0.0280 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.281
## Multiple R-squared:  0.506, Adjusted R-squared:  0.26
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
##  11 weights are ~= 1. The remaining 47 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    0.163  0.890  0.959   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          1.72e-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 1.459e+15  1      3.819e+07
## LastAuthorFemale  6.574e+14  1      2.564e+07
## Year              8.109e+30 13      1.545e+01
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.09e+00 -1.84e-01 -2.71e-16  2.10e-01  5.87e-01
```

```
##
```

```
## Coefficients:
```

```
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.6733     0.0890   7.56 2.3e-09 ***
## FirstAuthorFemale1  0.2167     0.0890   2.43 0.01923 *
## LastAuthorFemale1 -0.2494     0.1045  -2.39 0.02162 *
## Year1997           0.0134     0.1689   0.08 0.93725
## Year1998          -0.1021     0.1919  -0.53 0.59742
```

```

## Year2001          0.2557      0.0890      2.87  0.00635 **
## Year2002         -0.2913      0.0890     -3.27  0.00214 **
## Year2004          0.0497      0.0890      0.56  0.57931
## Year2005          0.5747      0.0890      6.46  8.8e-08 ***
## Year2006          0.5862      0.1371      4.27  0.00011 ***
## Year2007          0.2516      0.1409      1.79  0.08138 .
## Year2008          0.1116      0.1218      0.92  0.36472
## Year2009          0.3322      0.1080      3.08  0.00369 **
## Year2010          0.4190      0.1981      2.12  0.04039 *
## Year2011          0.4227      0.1316      3.21  0.00253 **
## Year2012          0.3566      0.1277      2.79  0.00786 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.38
## Multiple R-squared:  0.342, Adjusted R-squared:  0.106
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.390  0.900  0.963  0.927  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      1.72e-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 -3.651e+15  1      NaN
## Year              -3.651e+15 13      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,

```

```

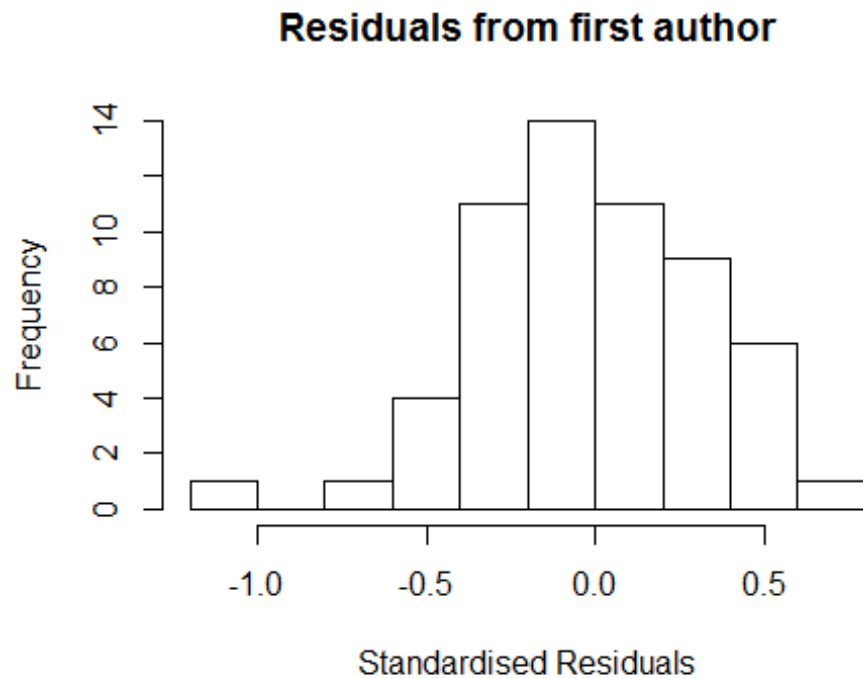
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q        Max
## -1.10e+00 -2.21e-01 -1.94e-16  2.31e-01  6.13e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.6947    0.0885   7.85 7.8e-10 ***
## FirstAuthorFemale1 0.1953    0.0885   2.21 0.03271 *
## Year1997         -0.2208    0.1388  -1.59 0.11917
## Year1998         -0.1128    0.1854  -0.61 0.54600
## Year2001          0.2343    0.0885   2.65 0.01129 *
## Year2002         -0.3127    0.0885  -3.53 0.00100 ***
## Year2004          0.0283    0.0885   0.32 0.75057
## Year2005          0.5533    0.0885   6.25 1.6e-07 ***
## Year2006          0.3153    0.0885   3.56 0.00091 ***
## Year2007          0.2331    0.1383   1.69 0.09907 .
## Year2008         -0.0238    0.0476  -0.50 0.61911
## Year2009          0.2845    0.1052   2.70 0.00978 **
## Year2010          0.4099    0.2115   1.94 0.05917 .
## Year2011          0.3742    0.1329   2.82 0.00732 **
## Year2012          0.3093    0.1243   2.49 0.01682 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.367
## Multiple R-squared:  0.32, Adjusted R-squared:  0.0984
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~ 1. The remaining 45 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.344  0.878  0.962   0.915  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      1.72e-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"

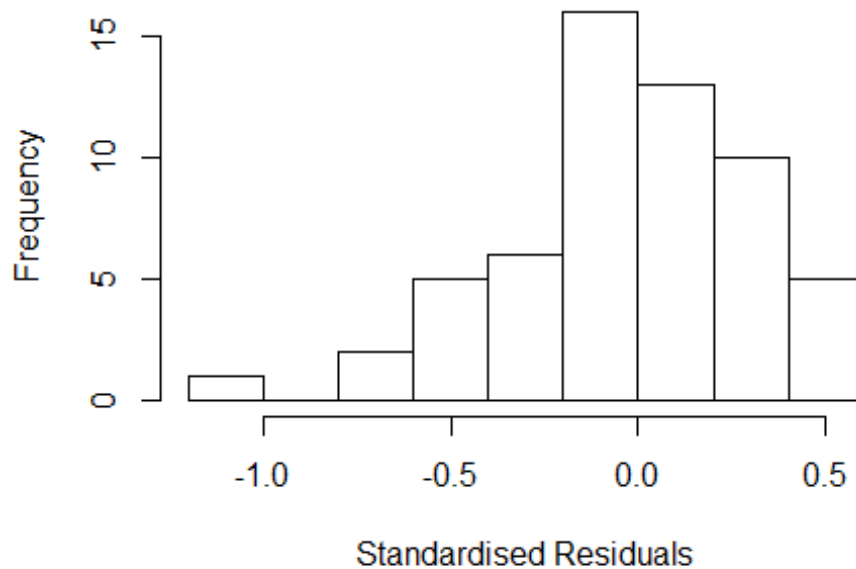
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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 13          NaN
```

Residuals from last author



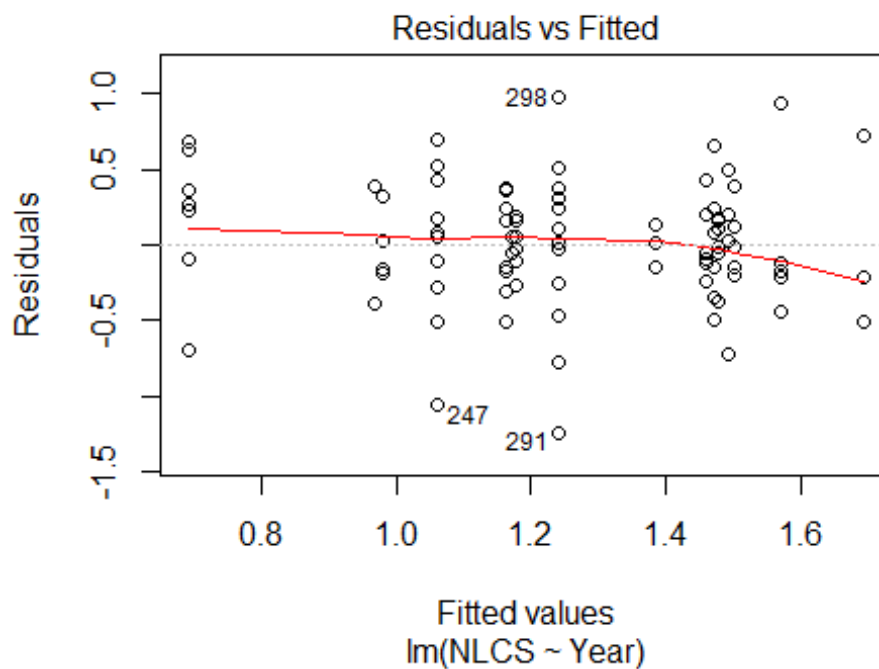
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.13e+00 -2.19e-01 2.88e-16 2.11e-01 5.66e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.90e-01 1.27e-08 7.03e+07 <2e-16 ***
## LastAuthorFemale1 -2.18e-01 1.11e-01 -1.96e+00 0.0570 .
## Year1997 -1.84e-01 1.66e-01 -1.11e+00 0.2731
## Year1998 -2.10e-01 9.90e-02 -2.13e+00 0.0392 *
## Year2001 3.90e-02 2.13e-08 1.83e+06 <2e-16 ***
## Year2002 -5.08e-01 1.93e-08 -2.63e+07 <2e-16 ***
## Year2004 -1.67e-01 0.00e+00 -Inf <2e-16 ***
## Year2005 3.58e-01 0.00e+00 Inf <2e-16 ***
## Year2006 3.38e-01 1.11e-01 3.03e+00 0.0041 **
## Year2007 6.47e-02 8.99e-02 7.20e-01 0.4757
## Year2008 -1.27e-02 1.85e-01 -7.00e-02 0.9454
## Year2009 1.68e-01 1.07e-01 1.56e+00 0.1252
```

```

## Year2010          2.39e-01  2.08e-01  1.15e+00  0.2553
## Year2011          3.58e-01  1.48e-01  2.42e+00  0.0197 *
## Year2012          1.61e-01  1.21e-01  1.32e+00  0.1923
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.307, Adjusted R-squared:  0.0817
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.339  0.895  0.964  0.920  0.986  0.996
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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:  58"
## [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
##    9    9   15   14    9    9    5   11    9    5    9   25   19   14   14
## 2011 2012
##   21   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    6    8    6    3    0    4    3    4    2    2    9    5    6    7
## 2011 2012
##   10   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```
##      5      6      8      3      3      0      4      3      3      2      2      8      5      5      6
## 2011 2012
##    10    10
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 26, df = 15, p-value = 0.04
```

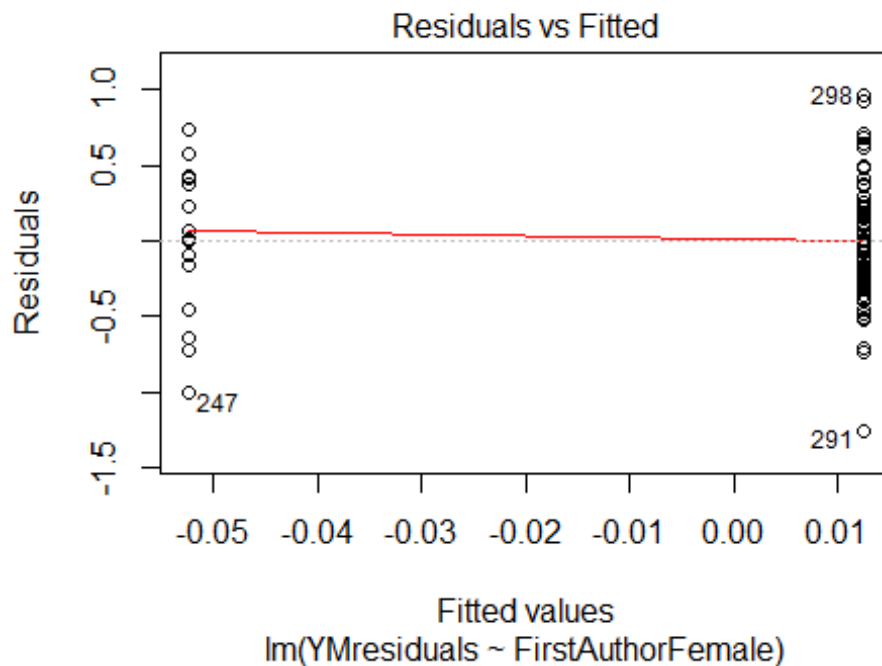


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.3, df = 1, p-value = 0.3
##
## [1] "Female first author team size 2018 geometric mean: 5"
## [1] "Male first author team size 2018 geometric mean: 4.61043629205845"
##
## Warning 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 = 1, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.16227766016838"
## [1] "Male last author team size 2018 geometric mean: 7"

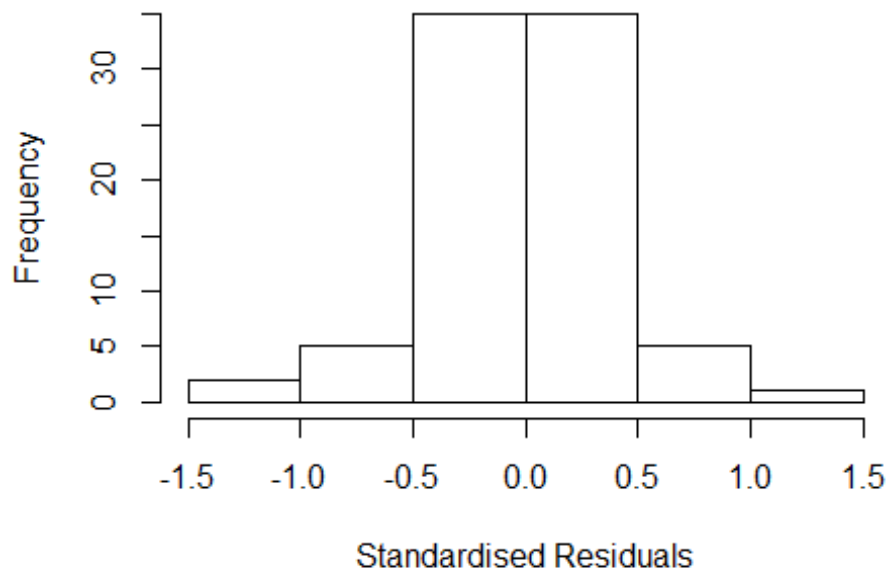
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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	9.393	1	3.065
LastAuthorFemale	5.688	1	2.385
UniqueAuthors	541.444	4	2.196
Year	1983.612	15	1.288

Residuals from first and last author and team size



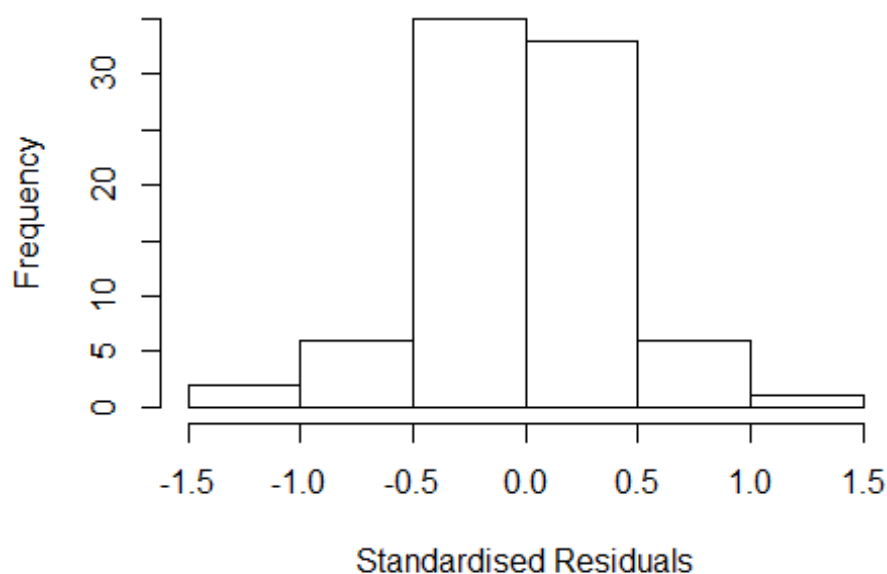
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2359 -0.2329 -0.0213 0.2596 1.0781
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6775 0.2986 5.62 5.1e-07 ***
## FirstAuthorFemale1 0.0511 0.1759 0.29 0.7724
## LastAuthorFemale1 -0.2181 0.1375 -1.59 0.1178
## UniqueAuthors2 0.0806 0.2896 0.28 0.7816
## UniqueAuthors3 -0.2041 0.2515 -0.81 0.4202
## UniqueAuthors4 -0.0967 0.3133 -0.31 0.7587
## UniqueAuthors5 -0.1355 0.2863 -0.47 0.6377
## Year1997 -0.1881 0.2972 -0.63 0.5292
## Year1998 -0.4046 0.2232 -1.81 0.0747 .
## Year1999 -0.1290 0.3030 -0.43 0.6719
```

```

## Year2000          -0.1713      0.2104    -0.81    0.4188
## Year2002          -0.1953      0.3271    -0.60    0.5528
## Year2003           0.2689      0.3140     0.86    0.3950
## Year2004          -0.5439      0.2089    -2.60    0.0116 *
## Year2005          -0.5388      0.3457    -1.56    0.1243
## Year2006          -0.5042      0.1935    -2.61    0.0115 *
## Year2007          -0.9809      0.3242    -3.03    0.0036 **
## Year2008          -0.0405      0.2487    -0.16    0.8711
## Year2009          -0.4092      0.1827    -2.24    0.0288 *
## Year2010          -0.1468      0.1655    -0.89    0.3785
## Year2011          -0.3925      0.2118    -1.85    0.0687 .
## Year2012          -0.4416      0.2536    -1.74    0.0867 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.386
## Multiple R-squared:  0.365, Adjusted R-squared:  0.146
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.285  0.885  0.951  0.904  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          1.20e-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 7.874 1          2.806
## LastAuthorFemale  4.660 1          2.159
## Year              17.182 15          1.099

```

Residuals from first and last author



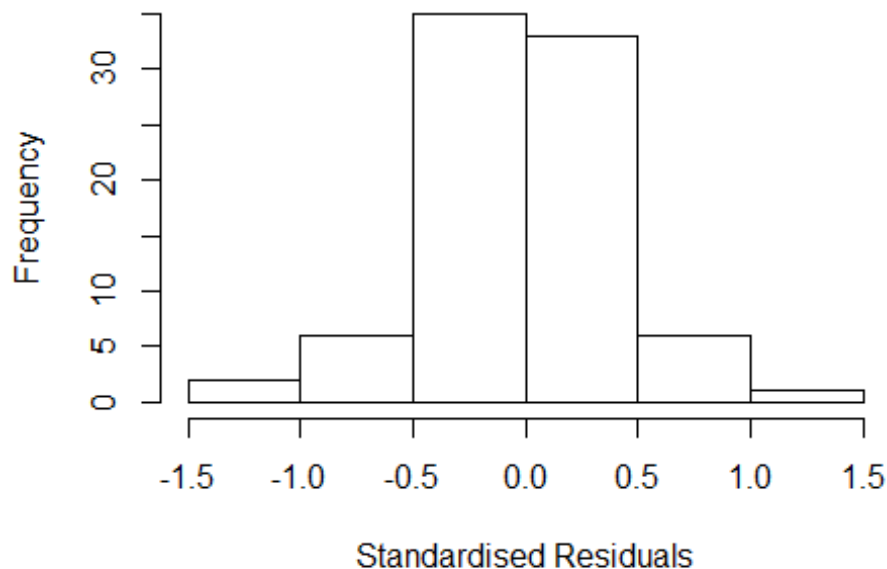
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2273 -0.2144 -0.0136 0.2551 1.0501
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5635 0.1141 13.70 <2e-16 ***
## FirstAuthorFemale1 0.0324 0.1560 0.21 0.8363
## LastAuthorFemale1 -0.1747 0.1322 -1.32 0.1910
## Year1997 -0.1134 0.2185 -0.52 0.6055
## Year1998 -0.4119 0.1699 -2.43 0.0181 *
## Year1999 -0.0149 0.1258 -0.12 0.9058
## Year2000 -0.1416 0.1699 -0.83 0.4077
## Year2002 -0.0208 0.2752 -0.08 0.9400
## Year2003 0.1979 0.3774 0.52 0.6018
## Year2004 -0.5577 0.2172 -2.57 0.0126 *
## Year2005 -0.5082 0.4792 -1.06 0.2929
## Year2006 -0.3218 0.1344 -2.39 0.0195 *
```

```

## Year2007          -0.8921      0.3171   -2.81    0.0065 **
## Year2008          -0.1026      0.2113   -0.49    0.6289
## Year2009          -0.3802      0.1384   -2.75    0.0078 **
## Year2010          -0.1165      0.1228   -0.95    0.3463
## Year2011          -0.3840      0.1849   -2.08    0.0418 *
## Year2012          -0.3362      0.2118   -1.59    0.1174
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.391
## Multiple R-squared:  0.311, Adjusted R-squared:  0.131
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 75 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.304  0.881  0.959  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      1.20e-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 5.736 1          2.395
## Year              5.736 15          1.060

```

Residuals from first author



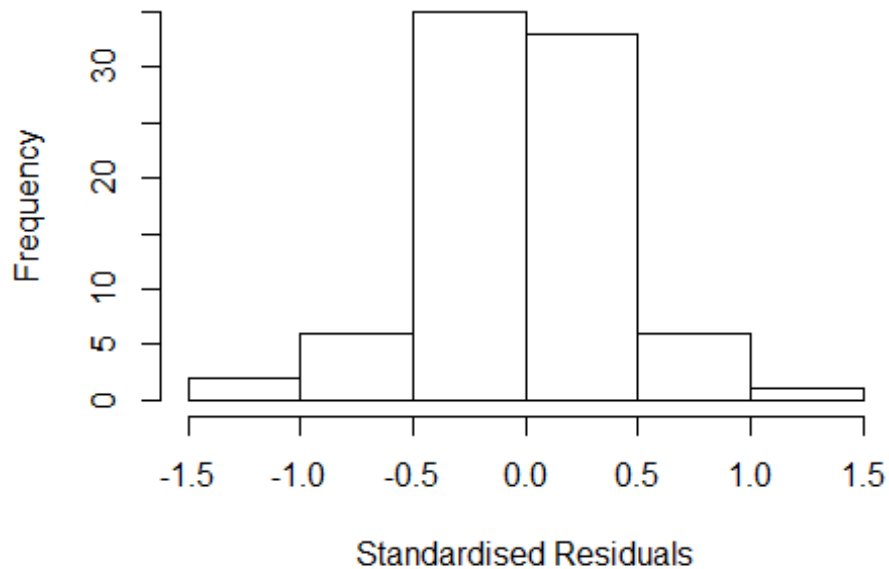
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2047 -0.1811 -0.0101 0.2586 1.0553
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5270 0.0966 15.81 <2e-16 ***
## FirstAuthorFemale1 -0.0142 0.1647 -0.09 0.9317
## Year1997 -0.0677 0.2064 -0.33 0.7440
## Year1998 -0.3539 0.1596 -2.22 0.0301 *
## Year1999 0.0216 0.1101 0.20 0.8452
## Year2000 -0.1336 0.1594 -0.84 0.4049
## Year2002 0.0182 0.2702 0.07 0.9464
## Year2003 0.0983 0.4546 0.22 0.8294
## Year2004 -0.5575 0.1764 -3.16 0.0024 **
## Year2005 -0.5590 0.3497 -1.60 0.1147
## Year2006 -0.3494 0.1303 -2.68 0.0093 **
## Year2007 -0.8394 0.3241 -2.59 0.0118 *
```

```

## Year2008          -0.0713      0.1996   -0.36   0.7221
## Year2009          -0.3436      0.1245   -2.76   0.0075 **
## Year2010          -0.1389      0.1103   -1.26   0.2124
## Year2011          -0.4049      0.1898   -2.13   0.0366 *
## Year2012          -0.3223      0.2169   -1.49   0.1421
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.295, Adjusted R-squared:  0.125
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.303  0.882  0.956  0.897  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      1.20e-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 3.843  1          1.960
## Year            3.843 15          1.046

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2335 -0.2110 -0.0174 0.2445 1.0516
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5615 0.1133 13.79 <2e-16 ***
## LastAuthorFemale1 -0.1650 0.1270 -1.30 0.1985
## Year1997 -0.1047 0.2065 -0.51 0.6137
## Year1998 -0.3951 0.1640 -2.41 0.0188 *
## Year1999 -0.0130 0.1250 -0.10 0.9177
## Year2000 -0.1216 0.1410 -0.86 0.3914
## Year2002 -0.0181 0.2755 -0.07 0.9478
## Year2003 0.2068 0.3665 0.56 0.5744
## Year2004 -0.5500 0.2243 -2.45 0.0169 *
## Year2005 -0.5110 0.4738 -1.08 0.2848
## Year2006 -0.3085 0.1179 -2.62 0.0110 *
## Year2007 -0.8794 0.3272 -2.69 0.0091 **
```

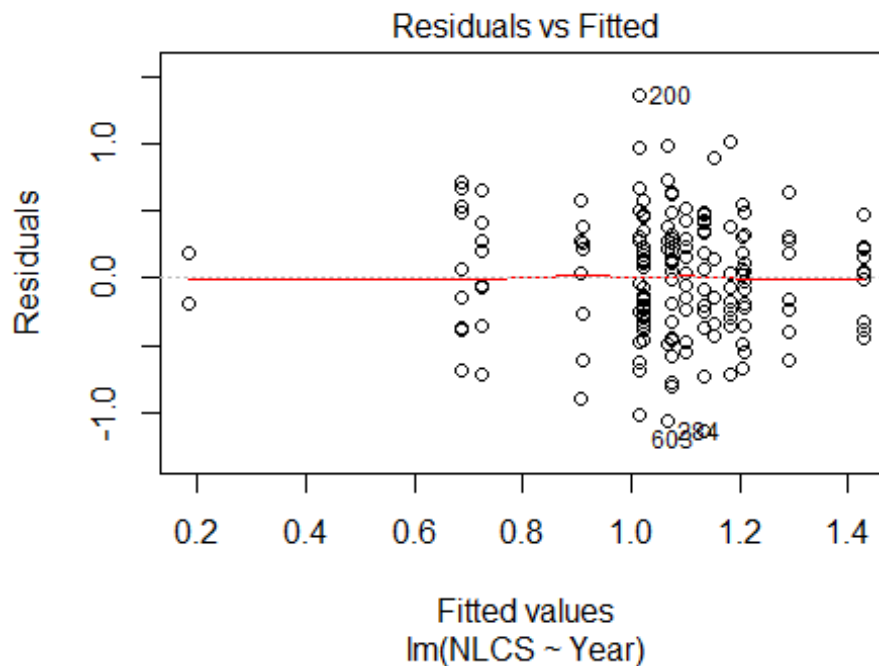


```

## Year2008          -0.1022      0.2096   -0.49    0.6275
## Year2009          -0.3782      0.1377   -2.75    0.0078 **
## Year2010          -0.1178      0.1215   -0.97    0.3358
## Year2011          -0.3797      0.1904   -1.99    0.0502 .
## Year2012          -0.3280      0.2116   -1.55    0.1259
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.389
## Multiple R-squared:  0.311, Adjusted R-squared:  0.144
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.293  0.876  0.960  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      1.20e-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: 83"
## [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
##   12   16   30   21   26   22   19   11   15   19   24   31   15   28   46
## 2011 2012
##   31   29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    8   12   10    7    9   13    4    6    8   10   12    5   13   20
## 2011 2012
##   23   14

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    7   11    6    6    7   11    4    5    8    9   11    5   12   16
## 2011 2012
##   19   10
## [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 = 0.29, df = 1, p-value = 0.6

## [1] "Female first author team size 2018 geometric mean: 4.58257569495584"
## [1] "Male first author team size 2018 geometric mean: 3.16814270910688"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

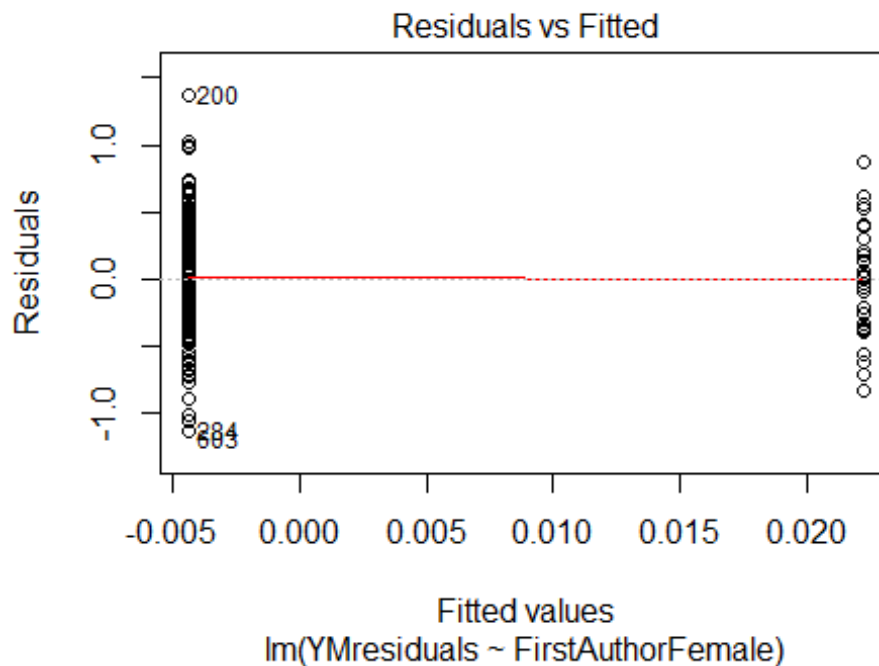
##
## Wilcoxon rank sum test with continuity correction
```

```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 3.32907107732884"

## 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 = 6.5, 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"

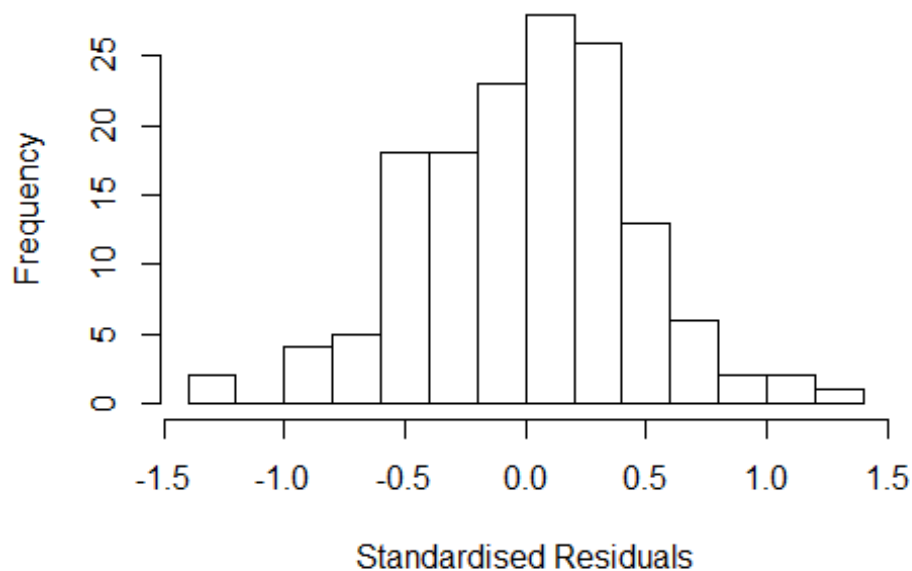
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```



```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.155 1      1.468
## LastAuthorFemale 2.424 1      1.557
```

## UniqueAuthors	56.499	4	1.656
## Year	85.818	16	1.149

Residuals from first and last author and team size



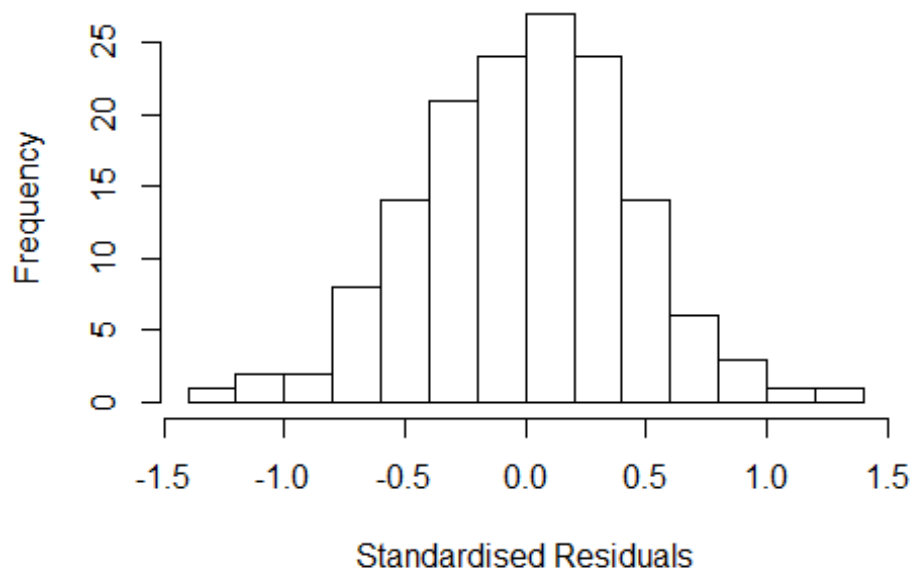
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2165 -0.3156 0.0212 0.2682 1.2380
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3640 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.0256 0.0997 -0.26 0.79781
## LastAuthorFemale1 0.0349 0.1029 0.34 0.73464
## UniqueAuthors2 0.1849 0.1353 1.37 0.17426
## UniqueAuthors3 0.1559 0.1262 1.23 0.21927
## UniqueAuthors4 0.2313 0.1582 1.46 0.14610
## UniqueAuthors5 0.2050 0.1660 1.23 0.21921
```

```

## Year1997      0.7135      0.1771      4.03  9.7e-05 ***
## Year1998      0.7290      0.1349      5.40  3.2e-07 ***
## Year1999      0.9899      0.1607      6.16  9.2e-09 ***
## Year2000      0.6152      0.1674      3.67  0.00035 ***
## Year2001      0.1530      0.2045      0.75  0.45577
## Year2002      0.6211      0.3634      1.71  0.08988 .
## Year2003      0.4156      0.3621      1.15  0.25320
## Year2004      0.6326      0.2285      2.77  0.00648 **
## Year2005      0.6290      0.3711      1.70  0.09255 .
## Year2006      0.5657      0.1730      3.27  0.00139 **
## Year2007      0.5424      0.1372      3.95  0.00013 ***
## Year2008      0.3965      0.1896      2.09  0.03859 *
## Year2009      0.1728      0.2024      0.85  0.39486
## Year2010      0.5869      0.1365      4.30  3.4e-05 ***
## Year2011      0.5054      0.1166      4.34  3.0e-05 ***
## Year2012      0.6845      0.1484      4.61  9.7e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.467
## Multiple R-squared:  0.162, Adjusted R-squared:  0.0145
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 132 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.463  0.887  0.953  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      6.76e-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 1.848 1      1.360
## LastAuthorFemale  1.780 1      1.334
## Year              2.954 16      1.034

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.208 -0.326 0.015 0.285 1.213
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 3.64e-01 8.73e-08 4.17e+06 < 2e-16 ***
## FirstAuthorFemale1 -1.28e-02 1.02e-01 -1.30e-01 0.89984
## LastAuthorFemale1 3.57e-02 9.99e-02 3.60e-01 0.72106
## Year1997 8.68e-01 1.60e-01 5.42e+00 2.8e-07 ***
## Year1998 8.51e-01 9.59e-02 8.87e+00 5.2e-15 ***
## Year1999 1.07e+00 1.42e-01 7.53e+00 7.8e-12 ***
## Year2000 7.59e-01 1.66e-01 4.56e+00 1.2e-05 ***
## Year2001 3.53e-01 1.70e-01 2.08e+00 0.03925 *
## Year2002 8.02e-01 2.62e-01 3.06e+00 0.00273 **
## Year2003 5.95e-01 3.11e-01 1.91e+00 0.05803 .
## Year2004 7.59e-01 2.22e-01 3.43e+00 0.00082 ***
## Year2005 8.08e-01 3.89e-01 2.08e+00 0.03974 *
```

```

## Year2006          7.21e-01    1.60e-01    4.50e+00    1.5e-05 ***
## Year2007          7.11e-01    1.06e-01    6.71e+00    5.7e-10 ***
## Year2008          5.62e-01    1.84e-01    3.06e+00    0.00269 **
## Year2009          3.50e-01    1.83e-01    1.91e+00    0.05873 .
## Year2010          7.34e-01    1.13e-01    6.52e+00    1.4e-09 ***
## Year2011          6.44e-01    7.59e-02    8.49e+00    4.3e-14 ***
## Year2012          8.60e-01    1.29e-01    6.68e+00    6.3e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.472
## Multiple R-squared:  0.138, Adjusted R-squared:  0.0173
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.489  0.898  0.955  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      6.76e-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 1.823 1      1.350
## Year              1.823 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

```

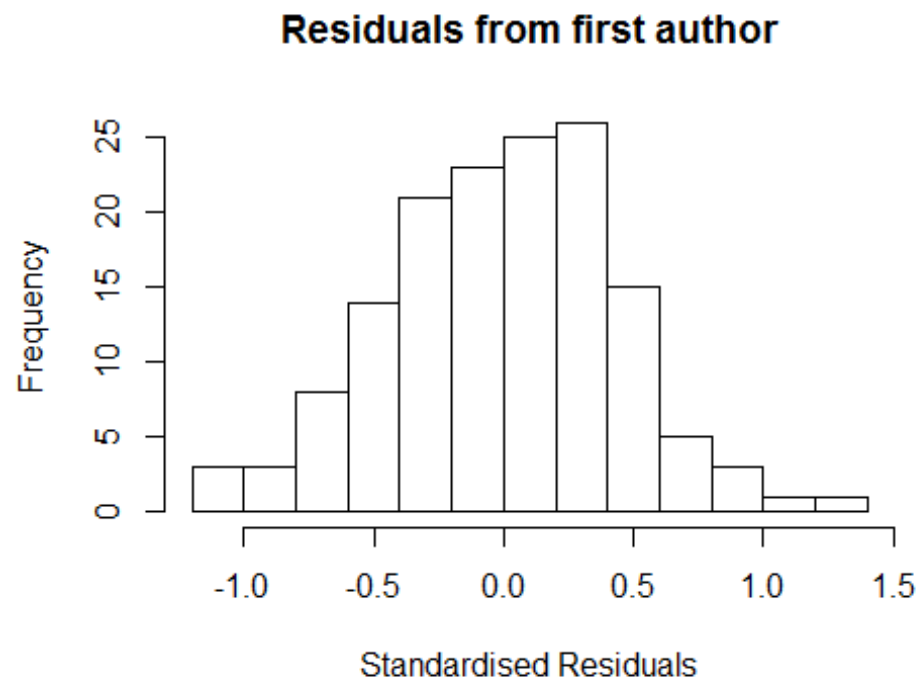
```

## -1.1789 -0.3328 0.0285 0.2792 1.2133
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    3.64e-01   1.27e-07  2.88e+06 < 2e-16 ***
## FirstAuthorFemale1 -4.61e-03   1.01e-01 -5.00e-02 0.96346
## Year1997        8.72e-01   1.62e-01  5.38e+00 3.3e-07 ***
## Year1998        8.48e-01   9.60e-02  8.83e+00 6.1e-15 ***
## Year1999        1.07e+00   1.42e-01  7.52e+00 7.8e-12 ***
## Year2000        7.63e-01   1.68e-01  4.55e+00 1.2e-05 ***
## Year2001        3.59e-01   1.66e-01  2.16e+00 0.03275 *
## Year2002        8.02e-01   2.64e-01  3.04e+00 0.00285 **
## Year2003        6.16e-01   3.07e-01  2.01e+00 0.04700 *
## Year2004        7.66e-01   2.19e-01  3.50e+00 0.00064 ***
## Year2005        8.15e-01   3.90e-01  2.09e+00 0.03850 *
## Year2006        7.28e-01   1.54e-01  4.74e+00 5.6e-06 ***
## Year2007        7.13e-01   1.08e-01  6.59e+00 1.0e-09 ***
## Year2008        5.62e-01   1.84e-01  3.06e+00 0.00270 **
## Year2009        3.55e-01   1.82e-01  1.94e+00 0.05406 .
## Year2010        7.40e-01   1.11e-01  6.70e+00 5.8e-10 ***
## Year2011        6.55e-01   7.08e-02  9.25e+00 5.7e-16 ***
## Year2012        8.61e-01   1.28e-01  6.71e+00 5.4e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.137, Adjusted R-squared:  0.0244
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 138 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.483 0.897 0.957 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          6.76e-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"

```

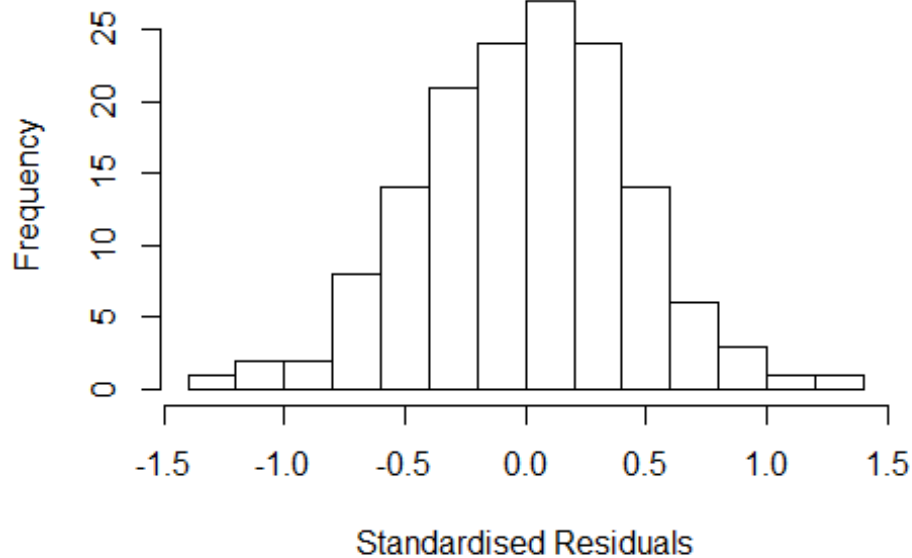


```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```



```
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.752 1          1.324
## Year             1.752 16          1.018
```

Residuals from last author



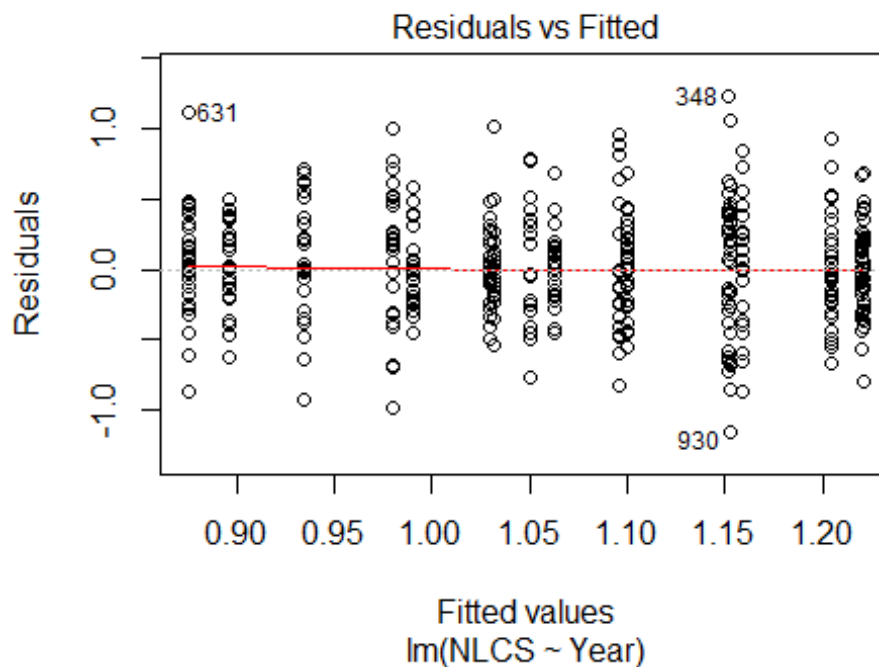
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2094 -0.3325 0.0151 0.2818 1.2134
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3640 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 0.0338 0.1001 0.34 0.73640
## Year1997 0.8665 0.1624 5.34 4.1e-07 ***
## Year1998 0.8466 0.0941 8.99 2.5e-15 ***
## Year1999 1.0673 0.1420 7.51 8.2e-12 ***
## Year2000 0.7555 0.1572 4.81 4.2e-06 ***
## Year2001 0.3540 0.1699 2.08 0.03921 *
## Year2002 0.8016 0.2654 3.02 0.00304 **
## Year2003 0.5955 0.3094 1.92 0.05642 .
## Year2004 0.7542 0.2232 3.38 0.00096 ***
## Year2005 0.8116 0.3860 2.10 0.03743 *
## Year2006 0.7157 0.1547 4.63 8.9e-06 ***
```

```

## Year2007          0.7102      0.1051      6.76  4.2e-10 ***
## Year2008          0.5625      0.1841      3.05  0.00273 **
## Year2009          0.3484      0.1870      1.86  0.06472 .
## Year2010          0.7327      0.1100      6.66  7.0e-10 ***
## Year2011          0.6441      0.0762      8.46  4.9e-14 ***
## Year2012          0.8561      0.1225      6.99  1.3e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.464
## Multiple R-squared:  0.138, Adjusted R-squared:  0.0258
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 138 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.474  0.896  0.956  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      6.76e-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: 148"
## [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
##   39   51   52   45   43   39   32   28   38   33   33   61   38   53   44
## 2011 2012
##   45   41
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   31   33   40   27   24   22   17   23   22   19   23   36   20   29   22
## 2011 2012

```

```
## 29 26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 27 32 36 24 21 19 13 21 17 16 18 30 16 24 13
## 2011 2012
## 25 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 = 62, df = 16, p-value = 2e-07
```



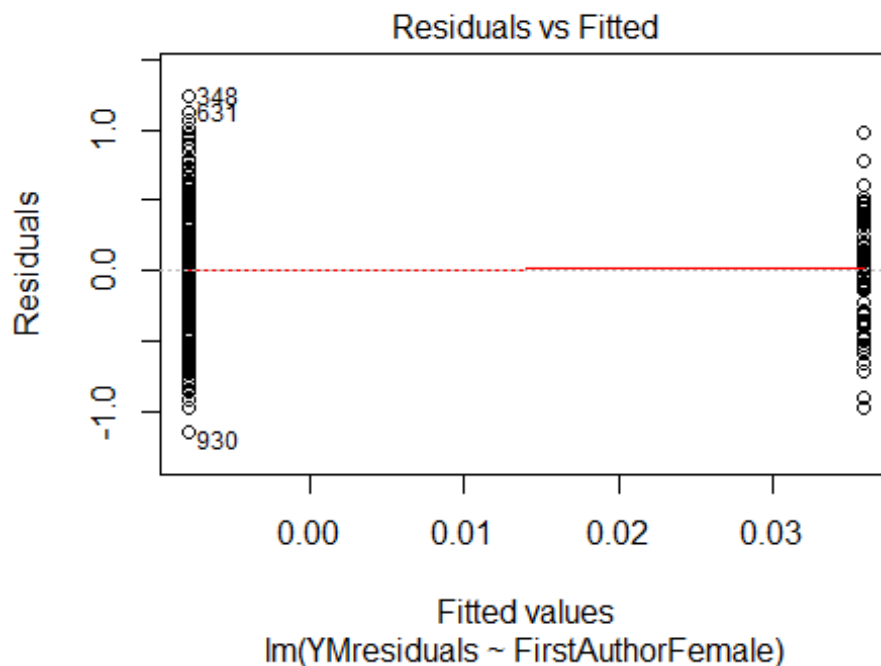
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.48, df = 1, p-value = 0.5

## [1] "Female first author team size 2018 geometric mean: 5"
## [1] "Male first author team size 2018 geometric mean: 4.61022289532816"

## 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 = 5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 4.61022289532816"

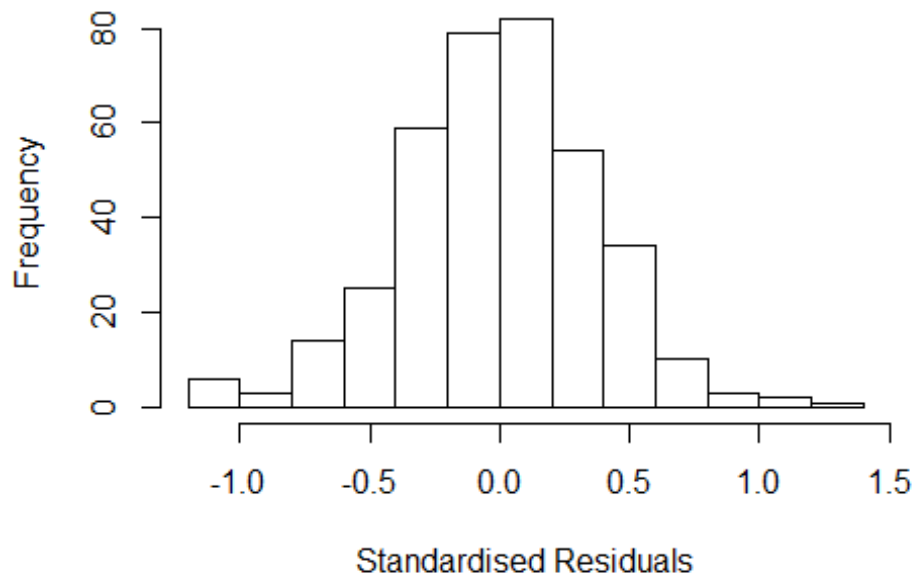
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5, 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	1.365	1	1.168
LastAuthorFemale	1.393	1	1.180
UniqueAuthors	2.406	4	1.116
Year	3.654	16	1.041

Residuals from first and last author and team size



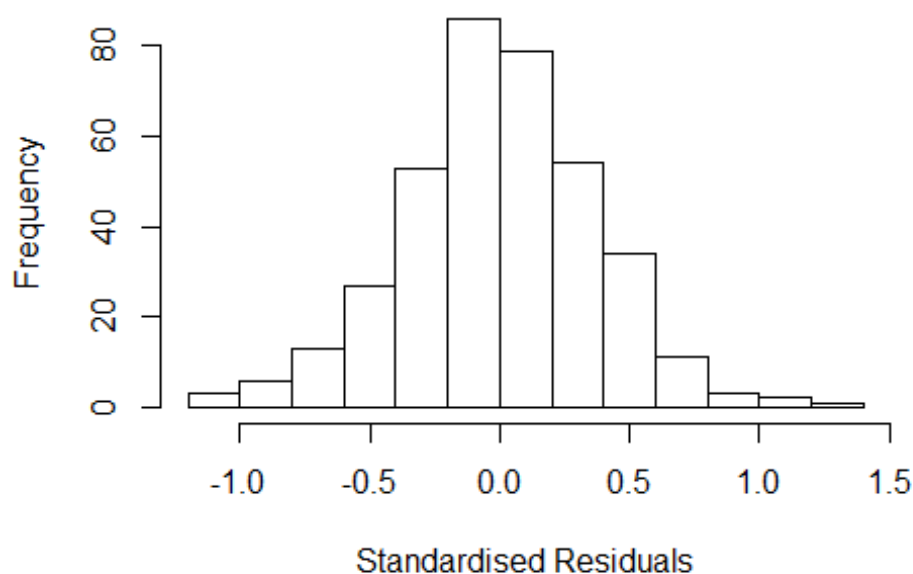
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.110417 -0.226008 0.000759 0.231395 1.226169
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.27197 0.14500 8.77 < 2e-16 ***
## FirstAuthorFemale1 0.05464 0.05381 1.02 0.31066
## LastAuthorFemale1 -0.04921 0.05911 -0.83 0.40564
## UniqueAuthors2 -0.00761 0.13750 -0.06 0.95587
## UniqueAuthors3 -0.03752 0.13574 -0.28 0.78239
## UniqueAuthors4 -0.02303 0.14330 -0.16 0.87241
## UniqueAuthors5 -0.09665 0.14381 -0.67 0.50201
## Year1997 -0.06692 0.08739 -0.77 0.44432
## Year1998 -0.14416 0.06866 -2.10 0.03649 *
## Year1999 -0.05499 0.09032 -0.61 0.54299
```

```

## Year2000      -0.22844    0.07485   -3.05  0.00245 **
## Year2001      -0.29343    0.08509   -3.45  0.00063 ***
## Year2002      -0.08162    0.16145   -0.51  0.61350
## Year2003      -0.23332    0.07026   -3.32  0.00099 ***
## Year2004      -0.26661    0.07302   -3.65  0.00030 ***
## Year2005      -0.18460    0.19133   -0.96  0.33530
## Year2006      -0.31178    0.08813   -3.54  0.00046 ***
## Year2007      -0.39152    0.09045   -4.33  2e-05 ***
## Year2008      -0.31875    0.10858   -2.94  0.00355 **
## Year2009      -0.24015    0.14083   -1.71  0.08904 .
## Year2010      -0.06422    0.10860   -0.59  0.55469
## Year2011      -0.19514    0.10969   -1.78  0.07611 .
## Year2012      -0.16156    0.16160   -1.00  0.31812
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.343
## Multiple R-squared:  0.097, Adjusted R-squared:  0.0401
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 341 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.176  0.860  0.952  0.893  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      2.69e-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 1.315 1      1.147
## LastAuthorFemale  1.330 1      1.153
## Year              1.683 16      1.016

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.08598 -0.21824 -0.00284 0.23279 1.20285
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2402 0.0516 24.05 < 2e-16 ***
## FirstAuthorFemale1 0.0506 0.0530 0.96 0.33993
## LastAuthorFemale1 -0.0479 0.0588 -0.81 0.41592
## Year1997 -0.0642 0.0851 -0.75 0.45122
## Year1998 -0.1419 0.0679 -2.09 0.03735 *
## Year1999 -0.0521 0.0875 -0.60 0.55185
## Year2000 -0.2295 0.0760 -3.02 0.00270 **
## Year2001 -0.2880 0.0827 -3.48 0.00056 ***
## Year2002 -0.0640 0.1614 -0.40 0.69194
## Year2003 -0.2324 0.0705 -3.29 0.00108 **
## Year2004 -0.2669 0.0751 -3.55 0.00043 ***
## Year2005 -0.1820 0.1967 -0.93 0.35539
```

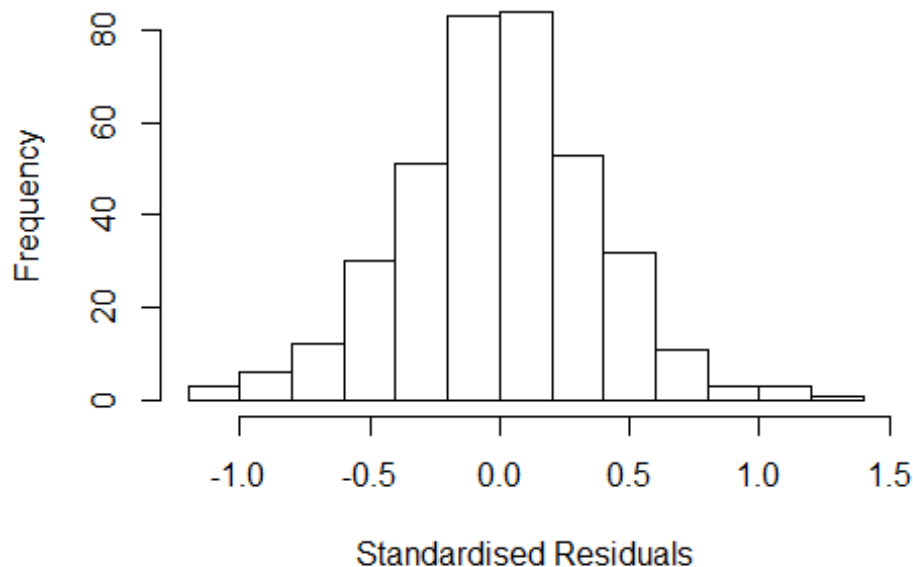


```

## Year2006          -0.3192      0.0867    -3.68  0.00027 ***
## Year2007          -0.3815      0.0911    -4.19  3.6e-05 ***
## Year2008          -0.3048      0.1055    -2.89  0.00411 **
## Year2009          -0.2443      0.1403    -1.74  0.08260 .
## Year2010          -0.0655      0.1045    -0.63  0.53110
## Year2011          -0.2048      0.1083    -1.89  0.05955 .
## Year2012          -0.1650      0.1627    -1.01  0.31141
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.0918, Adjusted R-squared:  0.0455
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 341 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.200  0.862  0.951  0.894  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      2.69e-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 1.3 1      1.140
## Year              1.3 16      1.008

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.08160 -0.21472 0.00148 0.23610 1.20643
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2286 0.0483 25.46 < 2e-16 ***
## FirstAuthorFemale1 0.0517 0.0527 0.98 0.32749
## Year1997 -0.0537 0.0838 -0.64 0.52172
## Year1998 -0.1345 0.0667 -2.02 0.04437 *
## Year1999 -0.0429 0.0857 -0.50 0.61726
## Year2000 -0.2247 0.0755 -2.98 0.00311 **
## Year2001 -0.2764 0.0808 -3.42 0.00070 ***
## Year2002 -0.0560 0.1576 -0.36 0.72232
## Year2003 -0.2249 0.0679 -3.31 0.00102 **
## Year2004 -0.2712 0.0724 -3.75 0.00021 ***
## Year2005 -0.1834 0.1971 -0.93 0.35294
## Year2006 -0.3203 0.0855 -3.75 0.00021 ***
```

```

## Year2007          -0.3746      0.0900   -4.16    4e-05 ***
## Year2008          -0.2972      0.1046   -2.84    0.00475 **
## Year2009          -0.2460      0.1419   -1.73    0.08378 .
## Year2010          -0.0714      0.1047   -0.68    0.49554
## Year2011          -0.1987      0.1086   -1.83    0.06810 .
## Year2012          -0.1630      0.1674   -0.97    0.33105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.347
## Multiple R-squared:  0.0906, Adjusted R-squared:  0.0469
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 340 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.201  0.864  0.952  0.894  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      2.69e-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 1.301 1      1.141
## Year      1.301 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.0877 -0.2285  0.0015  0.2269  1.2010

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2498    0.0493   25.36 < 2e-16 ***
## LastAuthorFemale1 -0.0496    0.0590   -0.84  0.40093
## Year1997          -0.0618    0.0839   -0.74  0.46186
## Year1998          -0.1416    0.0668   -2.12  0.03481 *
## Year1999          -0.0517    0.0866   -0.60  0.55089
## Year2000          -0.2240    0.0753   -2.98  0.00313 **
## Year2001          -0.2945    0.0812   -3.63  0.00033 ***
## Year2002          -0.0719    0.1572   -0.46  0.64789
## Year2003          -0.2259    0.0706   -3.20  0.00149 **
## Year2004          -0.2716    0.0750   -3.62  0.00034 ***
## Year2005          -0.1862    0.2034   -0.92  0.36051
## Year2006          -0.3142    0.0867   -3.62  0.00033 ***
## Year2007          -0.3904    0.0899   -4.34  1.8e-05 ***
## Year2008          -0.3073    0.1037   -2.96  0.00326 **
## Year2009          -0.2375    0.1376   -1.73  0.08513 .
## Year2010          -0.0614    0.1041   -0.59  0.55602
## Year2011          -0.2030    0.1086   -1.87  0.06234 .
## Year2012          -0.1622    0.1598   -1.02  0.31071
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.348
## Multiple R-squared:  0.0895, Adjusted R-squared:  0.0457
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 343 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.209  0.868  0.952  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           2.69e-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: 372"
## [1] ""

```

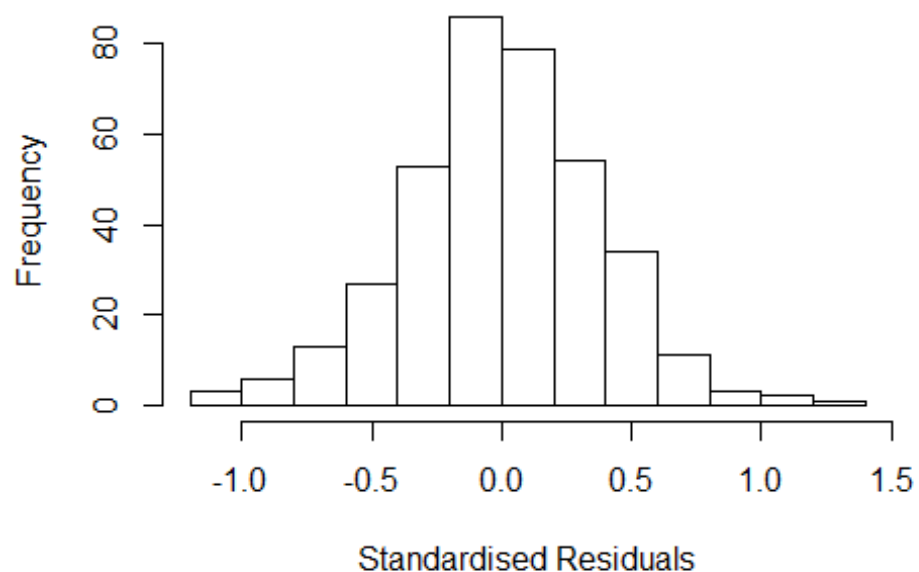
```

## [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
##      8   12   15   11    8   10   12    5    9    7    3   13    9    7   13
## 2011 2012
##     13   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      6    9    9    7    3    0    5    2    4    3    1    6    3    3    6
## 2011 2012
##      9   11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      5    8    8    4    3    0    5    2    3    3    1    5    3    3    4
## 2011 2012
##      9    9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5"
## [1] "Male first author team size 2018 geometric mean: 5.13451722960574"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

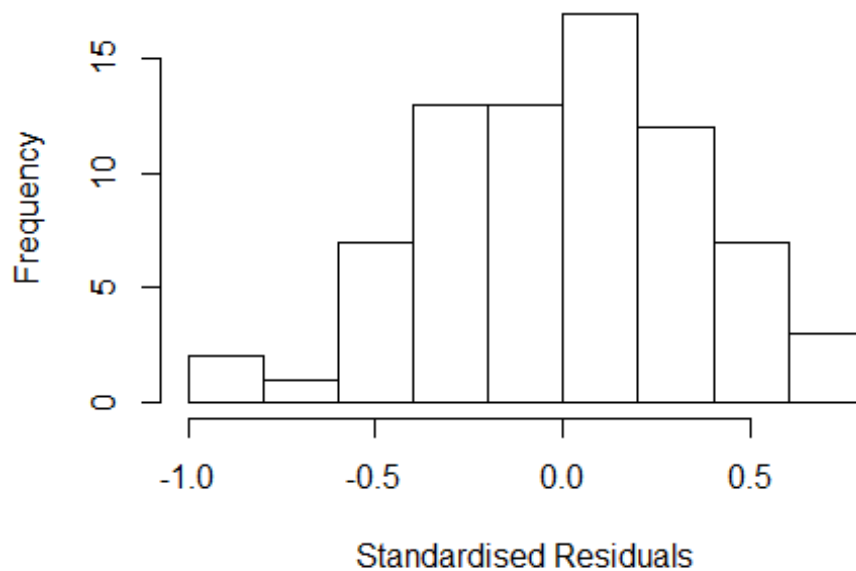
Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 5.10431513841152"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF ^{1/(2*Df)}
FirstAuthorFemale	17.34	1	4.165
LastAuthorFemale	14.92	1	3.863
UniqueAuthors	219.57	4	1.962
Year	1812.75	15	1.284

Residuals from first and last author and team size



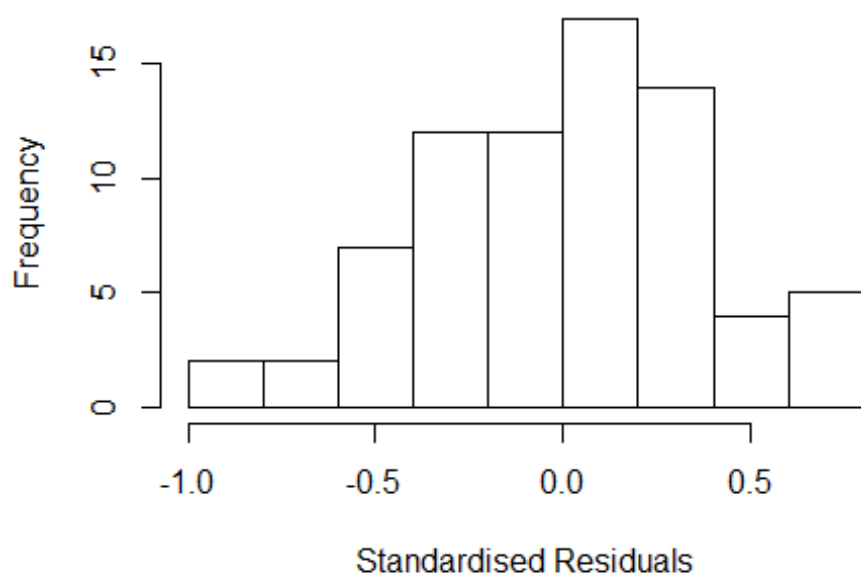
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9059 -0.2831 0.0121 0.2286 0.7696
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.66381 0.25700 6.47 3.2e-08 ***
## FirstAuthorFemale1 0.07929 0.15813 0.50 0.6182
## LastAuthorFemale1 0.02479 0.16766 0.15 0.8830
## UniqueAuthors2 -0.16531 0.20399 -0.81 0.4214
## UniqueAuthors3 -0.23282 0.22498 -1.03 0.3054
## UniqueAuthors4 -0.08656 0.24806 -0.35 0.7285
## UniqueAuthors5 0.00764 0.25808 0.03 0.9765
## Year1997 -0.04126 0.24632 -0.17 0.8676
## Year1998 -0.43422 0.18956 -2.29 0.0260 *
## Year1999 0.00852 0.28078 0.03 0.9759
```

```

## Year2000      -0.20809    0.21568   -0.96    0.3390
## Year2002      0.17843    0.34633    0.52    0.6086
## Year2003     -0.15994    0.19986   -0.80    0.4271
## Year2004     -0.21840    0.30624   -0.71    0.4789
## Year2005     -0.17782    0.47391   -0.38    0.7090
## Year2006     -0.48959    0.19468   -2.51    0.0150 *
## Year2007     -1.28670    0.25328   -5.08  5.0e-06 ***
## Year2008     -0.69965    0.24593   -2.84    0.0063 **
## Year2009     -0.70160    0.34656   -2.02    0.0480 *
## Year2010     -0.44592    0.23577   -1.89    0.0640 .
## Year2011     -0.52666    0.17797   -2.96    0.0046 **
## Year2012     -0.23838    0.19807   -1.20    0.2341
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.465, Adjusted R-squared:  0.253
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.610  0.897  0.955  0.926  0.988  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.33e-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 11.52 1      3.394
## LastAuthorFemale  12.24 1      3.498
## Year              54.35 15      1.142

```


Residuals from first and last author



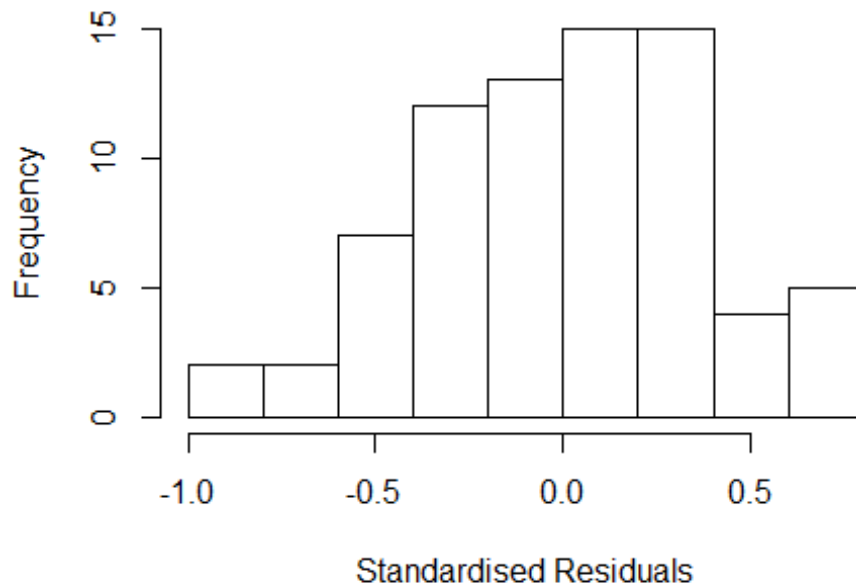
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9563 -0.2367 0.0157 0.2565 0.7050
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5398 0.1062 14.50 < 2e-16 ***
## FirstAuthorFemale1 0.1356 0.1260 1.08 0.28634
## LastAuthorFemale1 -0.0573 0.1480 -0.39 0.70025
## Year1997 -0.1118 0.2130 -0.52 0.60166
## Year1998 -0.4676 0.1533 -3.05 0.00347 **
## Year1999 0.0960 0.1376 0.70 0.48816
## Year2000 -0.2273 0.1660 -1.37 0.17645
## Year2002 0.1771 0.2879 0.62 0.54087
## Year2003 -0.2148 0.1587 -1.35 0.18106
## Year2004 -0.2405 0.3513 -0.68 0.49637
## Year2005 -0.1639 0.5479 -0.30 0.76595
## Year2006 -0.5052 0.1725 -2.93 0.00489 **
```

```

## Year2007          -1.2827      0.2441    -5.26  2.3e-06 ***
## Year2008          -0.6646      0.1868    -3.56  0.00076 ***
## Year2009          -0.7103      0.2682    -2.65  0.01043 *
## Year2010          -0.3706      0.2387    -1.55  0.12605
## Year2011          -0.5004      0.1469    -3.41  0.00121 **
## Year2012          -0.2385      0.1862    -1.28  0.20541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.428
## Multiple R-squared:  0.448, Adjusted R-squared:  0.284
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.596  0.906  0.967  0.928  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      1.33e-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 10.29 1      3.208
## Year              10.29 15      1.081

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.9540 -0.2260  0.0116  0.2483  0.7106
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5276     0.0962   15.88 < 2e-16 ***
## FirstAuthorFemale1  0.1257     0.1237    1.02  0.31383
## Year1997         -0.0989     0.2073   -0.48  0.63505
## Year1998         -0.4505     0.1439   -3.13  0.00273 **
## Year1999          0.1082     0.1301    0.83  0.40894
## Year2000         -0.2283     0.1654   -1.38  0.17277
## Year2002          0.1927     0.2860    0.67  0.50308
## Year2003         -0.2549     0.1296   -1.97  0.05391 .
## Year2004         -0.2501     0.3511   -0.71  0.47904
## Year2005         -0.1871     0.5159   -0.36  0.71812
## Year2006         -0.5403     0.1572   -3.44  0.00109 **
## Year2007         -1.2702     0.2387   -5.32  1.7e-06 ***
```

```

## Year2008          -0.6522      0.1820   -3.58  0.00069 ***
## Year2009          -0.6973      0.2663   -2.62  0.01126 *
## Year2010          -0.3899      0.2108   -1.85  0.06944 .
## Year2011          -0.5072      0.1454   -3.49  0.00093 ***
## Year2012          -0.2286      0.1861   -1.23  0.22420
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.417
## Multiple R-squared:  0.449, Adjusted R-squared:  0.297
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.579  0.892  0.966  0.924  0.982  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      1.33e-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 12.99  1          3.604
## Year              12.99 15          1.089

## [1] "List of 0 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
## -0.9911 -0.2460  0.0143  0.2508  0.6914
##

```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5343    0.1049   14.63 < 2e-16 ***
## LastAuthorFemale1 -0.0311    0.1469   -0.21  0.8330
## Year1997          -0.0847    0.2026   -0.42  0.6773
## Year1998          -0.3999    0.1516   -2.64  0.0107 *
## Year1999           0.1015    0.1366    0.74  0.4603
## Year2000          -0.1399    0.1207   -1.16  0.2511
## Year2002           0.1832    0.2880    0.64  0.5273
## Year2003          -0.1677    0.1807   -0.93  0.3573
## Year2004          -0.1530    0.4026   -0.38  0.7053
## Year2005          -0.1747    0.5455   -0.32  0.7499
## Year2006          -0.3902    0.1459   -2.67  0.0097 **
## Year2007          -1.2480    0.2260   -5.52  8.3e-07 ***
## Year2008          -0.6591    0.1862   -3.54  0.0008 ***
## Year2009          -0.7047    0.2680   -2.63  0.0109 *
## Year2010          -0.3800    0.2334   -1.63  0.1089
## Year2011          -0.4926    0.1541   -3.20  0.0023 **
## Year2012          -0.1982    0.1728   -1.15  0.2561
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.444, Adjusted R-squared:  0.291
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 66 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.567  0.892  0.962  0.925  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      1.33e-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: 75"
## [1] ""
## [1] ""
## [1] "#####"

```

```

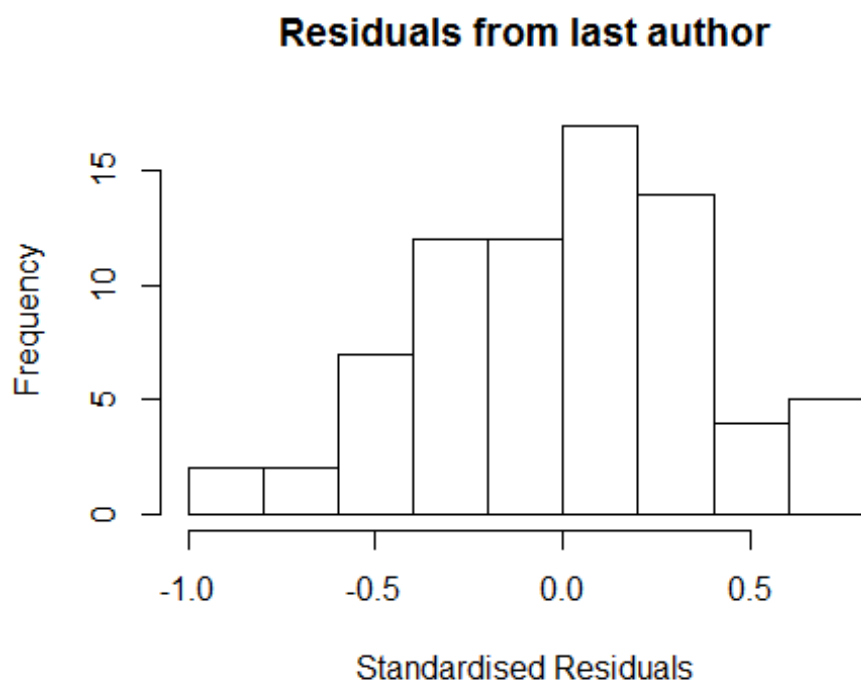
## [1] "Analysis of AJSC 2507"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 18 30 23 11 15 22 17 13 10 17 17 23 18 26
## 2011 2012
## 19 16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 6 4 7 1 4 5 8 4 5 8 8 12 11 12
## 2011 2012
## 12 9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 6 2 5 0 4 5 8 3 5 4 6 11 9 9
## 2011 2012
## 12 5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.82847668325146"
## [1] "Male first author team size 2018 geometric mean: 3.36358566101486"

## 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 = 9, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.88449914061482"
## [1] "Male last author team size 2018 geometric mean: 5.7005397655436"

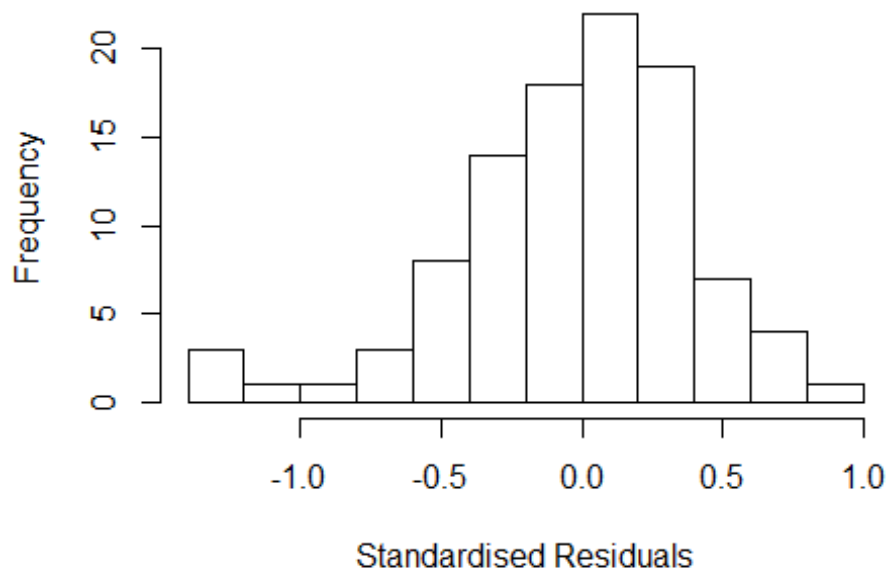
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, 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 2.063 1      1.436
## LastAuthorFemale  1.884 1      1.373
## UniqueAuthors    39.340 4      1.583
## Year              48.314 15     1.138
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.2324 -0.2508 0.0289 0.2674 0.8124
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.85576 0.18048 4.74 9.2e-06 ***
## FirstAuthorFemale1 0.07111 0.09268 0.77 0.445
## LastAuthorFemale1 0.02310 0.14927 0.15 0.877
## UniqueAuthors2 0.10499 0.17913 0.59 0.559
## UniqueAuthors3 -0.06573 0.20386 -0.32 0.748
## UniqueAuthors4 0.08296 0.19144 0.43 0.666
## UniqueAuthors5 -0.00751 0.24197 -0.03 0.975
## Year1997 0.16648 0.26371 0.63 0.530
## Year1998 0.05201 0.33262 0.16 0.876
## Year1999 -0.60497 0.27110 -2.23 0.028 *
```

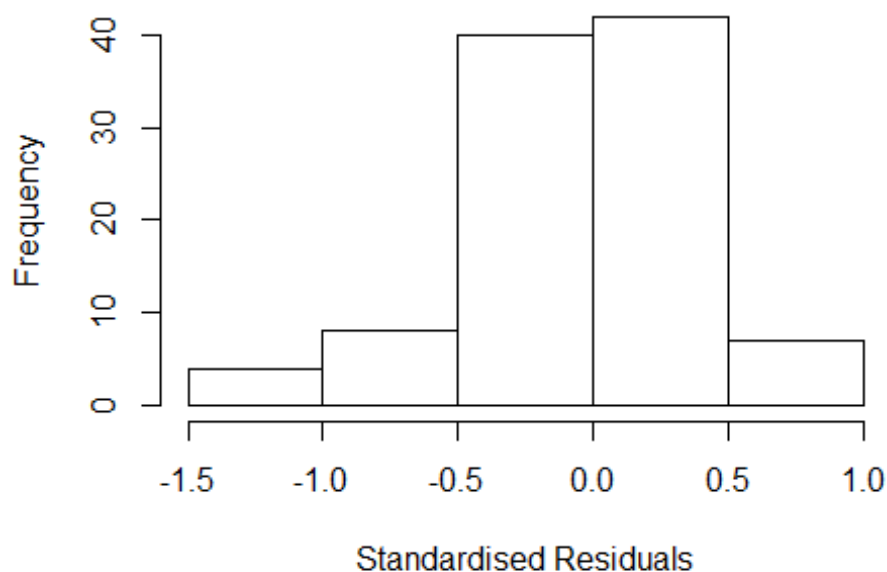


```

## Year2001      -0.07500    0.16078   -0.47    0.642
## Year2002      0.35145    0.37398    0.94    0.350
## Year2003     -0.38116    0.26820   -1.42    0.159
## Year2004      0.63253    0.29991    2.11    0.038 *
## Year2005      0.18241    0.17362    1.05    0.297
## Year2006      0.36799    0.17994    2.05    0.044 *
## Year2007      0.49172    0.31314    1.57    0.120
## Year2008      0.31639    0.15804    2.00    0.049 *
## Year2009      0.35454    0.22169    1.60    0.114
## Year2010      0.37660    0.21866    1.72    0.089 .
## Year2011      0.21486    0.21399    1.00    0.318
## Year2012      0.53643    0.22380    2.40    0.019 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.432
## Multiple R-squared:  0.374, Adjusted R-squared:  0.207
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 93 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.395  0.892  0.958  0.915  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      9.90e-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 1.952 1      1.397
## LastAuthorFemale  1.416 1      1.190
## Year              2.696 15      1.034

```

Residuals from first and last author



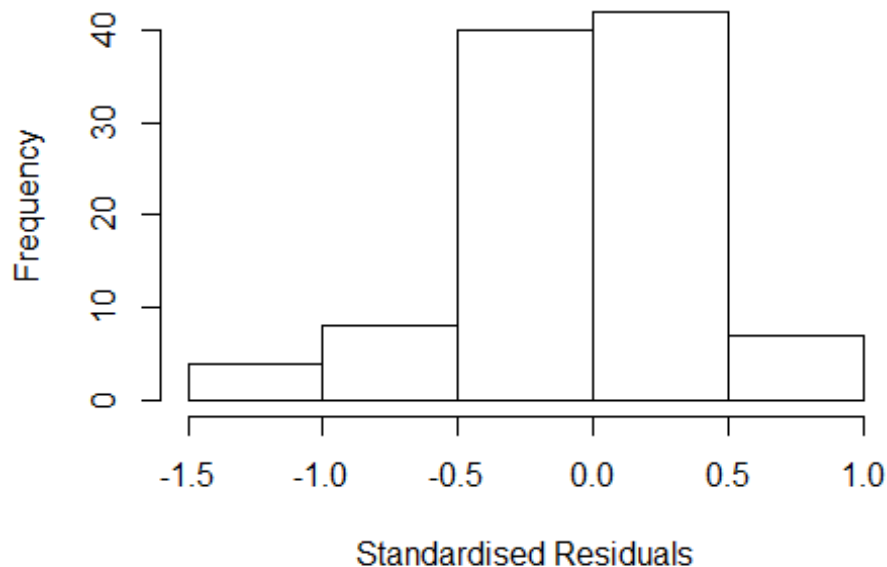
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3630 -0.2549 -0.0187 0.2508 0.9121
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8712 0.1558 5.59 2.8e-07 ***
## FirstAuthorFemale1 0.0781 0.0899 0.87 0.387
## LastAuthorFemale1 -0.0274 0.1366 -0.20 0.841
## Year1997 0.1973 0.2717 0.73 0.470
## Year1998 0.0780 0.5260 0.15 0.882
## Year1999 -0.5916 0.2626 -2.25 0.027 *
## Year2001 -0.0535 0.1639 -0.33 0.745
## Year2002 0.4156 0.3412 1.22 0.227
## Year2003 -0.3913 0.2877 -1.36 0.177
## Year2004 0.6684 0.3635 1.84 0.070 .
## Year2005 0.1729 0.1940 0.89 0.375
## Year2006 0.4160 0.1618 2.57 0.012 *
```

```

## Year2007          0.4920      0.3081      1.60      0.114
## Year2008          0.3562      0.1631      2.18      0.032 *
## Year2009          0.4138      0.1837      2.25      0.027 *
## Year2010          0.3978      0.1626      2.45      0.017 *
## Year2011          0.2374      0.1847      1.29      0.202
## Year2012          0.5780      0.2272      2.54      0.013 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.378
## Multiple R-squared:  0.406, Adjusted R-squared:  0.285
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 91 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.166  0.858  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      9.90e-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 1.991 1      1.411
## Year              1.991 15      1.023

```

Residuals from first author



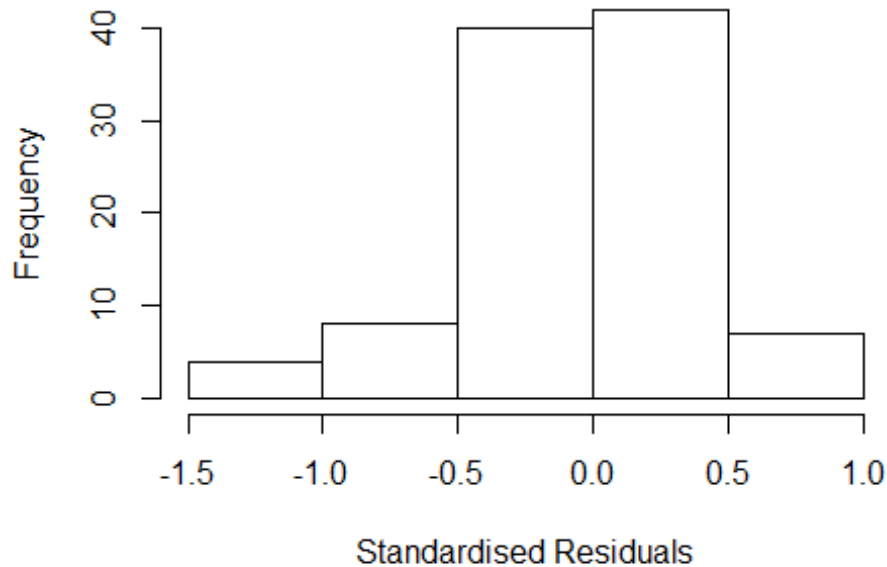
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3575 -0.2488 -0.0192 0.2544 0.9086
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8660 0.1490 5.81 1.1e-07 ***
## FirstAuthorFemale1 0.0786 0.0907 0.87 0.3885
## Year1997 0.1892 0.2791 0.68 0.4998
## Year1998 0.0692 0.4852 0.14 0.8869
## Year1999 -0.5842 0.2563 -2.28 0.0252 *
## Year2001 -0.0484 0.1572 -0.31 0.7590
## Year2002 0.4076 0.3424 1.19 0.2373
## Year2003 -0.3825 0.2792 -1.37 0.1743
## Year2004 0.6680 0.3588 1.86 0.0662 .
## Year2005 0.1715 0.1935 0.89 0.3780
## Year2006 0.4210 0.1552 2.71 0.0081 **
## Year2007 0.4936 0.2932 1.68 0.0960 .
```

```

## Year2008          0.3613      0.1565      2.31      0.0234 *
## Year2009          0.4129      0.1807      2.29      0.0248 *
## Year2010          0.3931      0.1631      2.41      0.0181 *
## Year2011          0.2388      0.1827      1.31      0.1946
## Year2012          0.5765      0.2277      2.53      0.0132 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.396, Adjusted R-squared:  0.281
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 90 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.200  0.867   0.953   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      9.90e-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 1.41  1          1.187
## Year              1.41 15          1.012

```

Residuals from last author



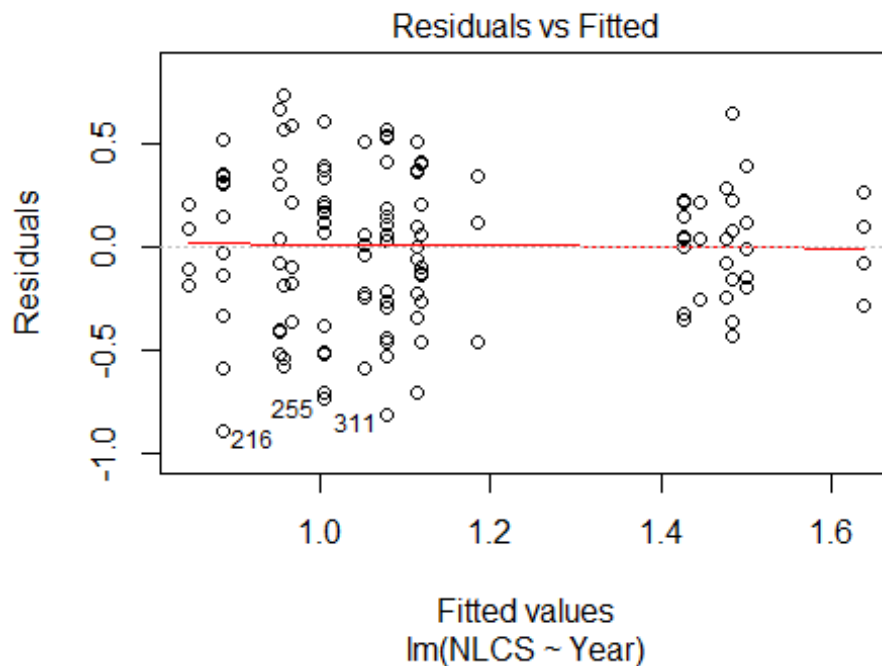
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.326 -0.263 -0.011  0.215  0.915
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8969    0.1498   5.99 5.1e-08 ***
## LastAuthorFemale1 -0.0176    0.1381  -0.13  0.8988
## Year1997        0.2121    0.2633   0.81  0.4227
## Year1998        0.0864    0.4543   0.19  0.8496
## Year1999       -0.6188    0.2600  -2.38  0.0196 *
## Year2001       -0.0590    0.1661  -0.36  0.7232
## Year2002        0.3986    0.3376   1.18  0.2411
## Year2003       -0.4194    0.2870  -1.46  0.1477
## Year2004        0.6465    0.3636   1.78  0.0791 .
## Year2005        0.1455    0.1907   0.76  0.4477
## Year2006        0.4093    0.1689   2.42  0.0175 *
## Year2007        0.5168    0.3303   1.56  0.1215
```

```

## Year2008          0.3721      0.1710      2.18      0.0324 *
## Year2009          0.4171      0.1825      2.29      0.0248 *
## Year2010          0.4293      0.1574      2.73      0.0078 **
## Year2011          0.2370      0.1890      1.25      0.2134
## Year2012          0.5567      0.2218      2.51      0.0140 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.411, Adjusted R-squared:  0.299
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 94 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.174  0.857  0.954  0.886  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      9.90e-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: 101"
## [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
##    8   13   11   16   12   11   13    6   11    9   19   18   10   28   29
## 2011 2012
##   23   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    6    9    8    4    4    5    3    3    4    8    8    6   11   15
## 2011 2012
##   17    9

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6    8    3    3    4    4    3    2    4    7    7    4    9    13
## 2011 2012
##   15    7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```

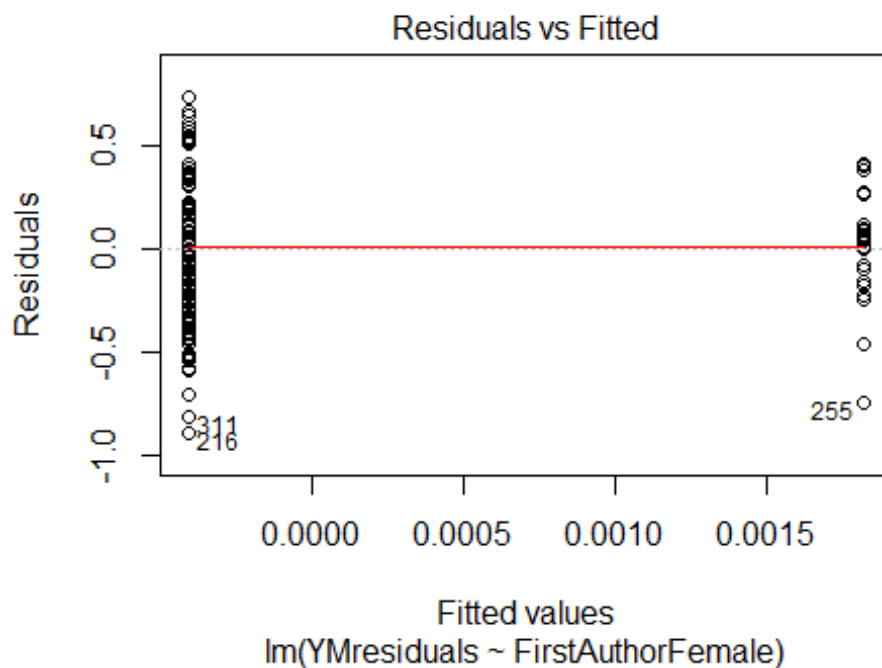


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.4, df = 1, p-value = 0.07
## [1] "Female first author team size 2018 geometric mean: 3.76447449532735"
## [1] "Male first author team size 2018 geometric mean: 3.95685131143309"
## 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 = 22, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 4.153248768869"

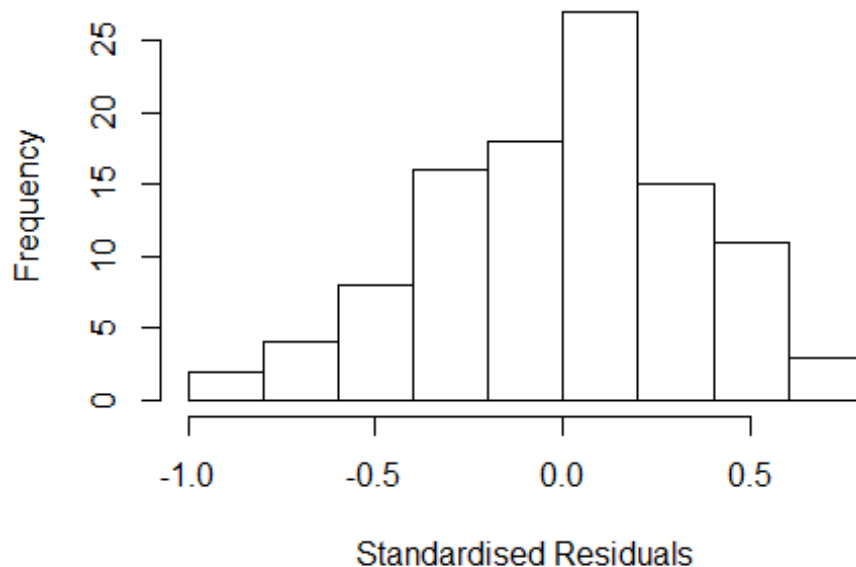
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 9, 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	3.215	1	1.793
LastAuthorFemale	2.602	1	1.613
UniqueAuthors	96.212	4	1.770
Year	252.168	16	1.189

Residuals from first and last author and team size



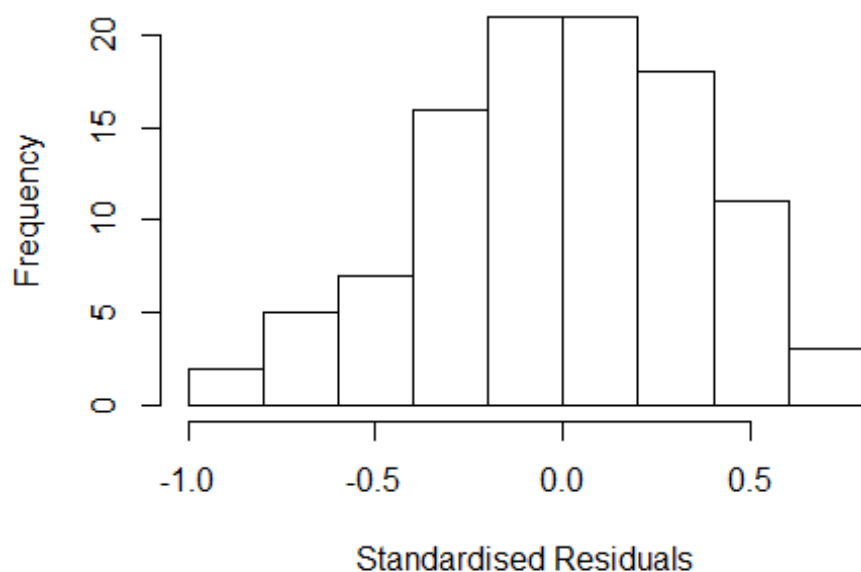
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9820 -0.2284 0.0223 0.2256 0.6959
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5886 0.2249 7.07 5.0e-10 ***
## FirstAuthorFemale1 0.0904 0.0859 1.05 0.29549
## LastAuthorFemale1 0.0535 0.0936 0.57 0.56939
## UniqueAuthors2 -0.0819 0.2258 -0.36 0.71783
## UniqueAuthors3 -0.0574 0.2068 -0.28 0.78217
## UniqueAuthors4 -0.0470 0.2288 -0.21 0.83787
## UniqueAuthors5 -0.1202 0.2486 -0.48 0.62995
## Year1997 -0.0696 0.1993 -0.35 0.72772
## Year1998 -0.4528 0.1430 -3.17 0.00217 **
## Year1999 -0.0400 0.2309 -0.17 0.86278
```

```

## Year2000          -0.1979      0.1435    -1.38    0.17156
## Year2001          -0.6978      0.1411    -4.95    4.0e-06 ***
## Year2002          -0.3905      0.4621    -0.84    0.40060
## Year2003          -0.1539      0.1846    -0.83    0.40675
## Year2004          -0.5667      0.2050    -2.76    0.00707 **
## Year2005           0.0292      0.1595     0.18    0.85505
## Year2006          -0.5914      0.1653    -3.58    0.00059 ***
## Year2007          -0.6696      0.1888    -3.55    0.00065 ***
## Year2008          -0.6532      0.1547    -4.22    6.3e-05 ***
## Year2009          -0.6065      0.2174    -2.79    0.00656 **
## Year2010          -0.4224      0.1359    -3.11    0.00260 **
## Year2011          -0.4509      0.1763    -2.56    0.01241 *
## Year2012          -0.4470      0.1896    -2.36    0.02079 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.38
## Multiple R-squared:  0.298, Adjusted R-squared:  0.107
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 94 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.484  0.891  0.953   0.920  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          9.62e-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 3.201 1          1.789
## LastAuthorFemale  1.909 1          1.382
## Year              4.896 16          1.051

```

Residuals from first and last author



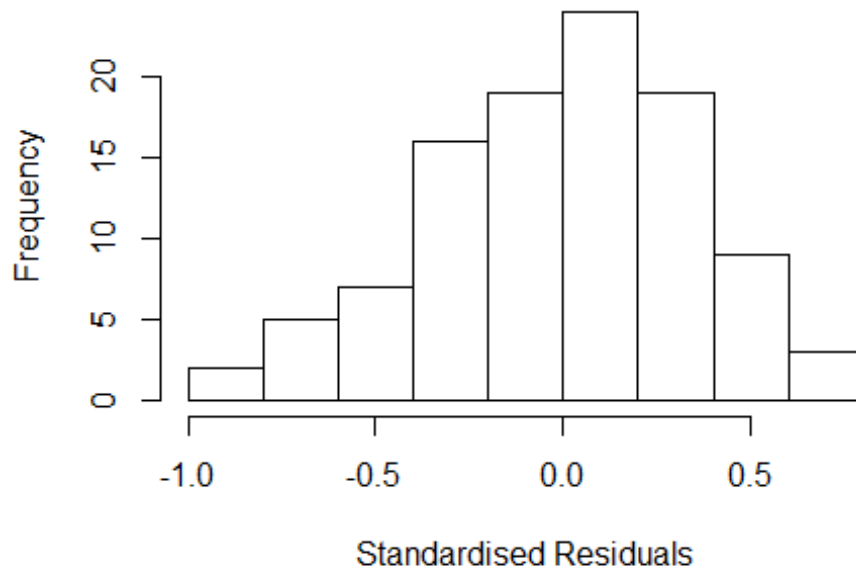
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8962 -0.2213 0.0107 0.2538 0.7129
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5160 0.0977 15.51 < 2e-16 ***
## FirstAuthorFemale1 0.0859 0.0866 0.99 0.32382
## LastAuthorFemale1 0.0498 0.0833 0.60 0.55136
## Year1997 -0.0689 0.2009 -0.34 0.73228
## Year1998 -0.4470 0.1418 -3.15 0.00224 **
## Year1999 0.0326 0.1111 0.29 0.77006
## Year2000 -0.2073 0.1468 -1.41 0.16169
## Year2001 -0.6717 0.1248 -5.38 6.4e-07 ***
## Year2002 -0.3970 0.4376 -0.91 0.36689
## Year2003 -0.1346 0.1728 -0.78 0.43829
## Year2004 -0.5719 0.1972 -2.90 0.00474 **
## Year2005 0.0317 0.1545 0.20 0.83813
```

```

## Year2006          -0.5887      0.1549    -3.80  0.00027 ***
## Year2007          -0.6587      0.1887    -3.49  0.00077 ***
## Year2008          -0.6554      0.1445    -4.54  1.9e-05 ***
## Year2009          -0.6198      0.2120    -2.92  0.00443 **
## Year2010          -0.4031      0.1407    -2.86  0.00526 **
## Year2011          -0.4272      0.1675    -2.55  0.01252 *
## Year2012          -0.4421      0.2015    -2.19  0.03094 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.296, Adjusted R-squared:  0.147
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 94 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.543  0.887  0.955  0.918  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      9.62e-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 3.049 1      1.746
## Year              3.049 16      1.035

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.90504 -0.22414 0.00843 0.25060 0.75246
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5267 0.0967 15.78 < 2e-16 ***
## FirstAuthorFemale1 0.0918 0.0863 1.06 0.29014
## Year1997 -0.0812 0.2002 -0.41 0.68614
## Year1998 -0.4603 0.1412 -3.26 0.00160 **
## Year1999 0.0219 0.1102 0.20 0.84324
## Year2000 -0.2046 0.1470 -1.39 0.16754
## Year2001 -0.6825 0.1241 -5.50 3.8e-07 ***
## Year2002 -0.4072 0.4395 -0.93 0.35681
## Year2003 -0.1132 0.1583 -0.72 0.47641
## Year2004 -0.5606 0.2147 -2.61 0.01063 *
## Year2005 0.0416 0.1497 0.28 0.78191
## Year2006 -0.5846 0.1572 -3.72 0.00036 ***
```

```

## Year2007          -0.6592      0.1961   -3.36   0.00116 **
## Year2008          -0.6661      0.1438   -4.63   1.3e-05 ***
## Year2009          -0.6217      0.2096   -2.97   0.00390 **
## Year2010          -0.4055      0.1434   -2.83   0.00584 **
## Year2011          -0.4194      0.1661   -2.53   0.01338 *
## Year2012          -0.4464      0.2034   -2.20   0.03085 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.295, Adjusted R-squared:  0.156
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 92 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.532  0.875   0.958   0.915   0.987   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      9.62e-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 1.738 1          1.318
## Year              1.738 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
## -0.8937 -0.2340  0.0142  0.2499  0.7056

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5142    0.0974   15.55 < 2e-16 ***
## LastAuthorFemale1  0.0585    0.0843    0.69  0.48963
## Year1997         -0.0487    0.1919   -0.25  0.80022
## Year1998         -0.4071    0.1491   -2.73  0.00765 **
## Year1999          0.0343    0.1108    0.31  0.75732
## Year2000         -0.1503    0.1161   -1.30  0.19872
## Year2001         -0.6700    0.1245   -5.38  6.3e-07 ***
## Year2002         -0.3960    0.4318   -0.92  0.36163
## Year2003         -0.1073    0.1670   -0.64  0.52230
## Year2004         -0.5314    0.2224   -2.39  0.01906 *
## Year2005          0.0952    0.1481    0.64  0.52172
## Year2006         -0.5606    0.1504   -3.73  0.00035 ***
## Year2007         -0.6583    0.1872   -3.52  0.00070 ***
## Year2008         -0.6535    0.1442   -4.53  1.9e-05 ***
## Year2009         -0.6205    0.2100   -2.95  0.00403 **
## Year2010         -0.3877    0.1378   -2.81  0.00607 **
## Year2011         -0.4290    0.1666   -2.57  0.01175 *
## Year2012         -0.3955    0.1722   -2.30  0.02406 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.377
## Multiple R-squared:  0.292, Adjusted R-squared:  0.152
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 92 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.553  0.884  0.953  0.917  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      9.62e-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: 104"
## [1] ""

```



```

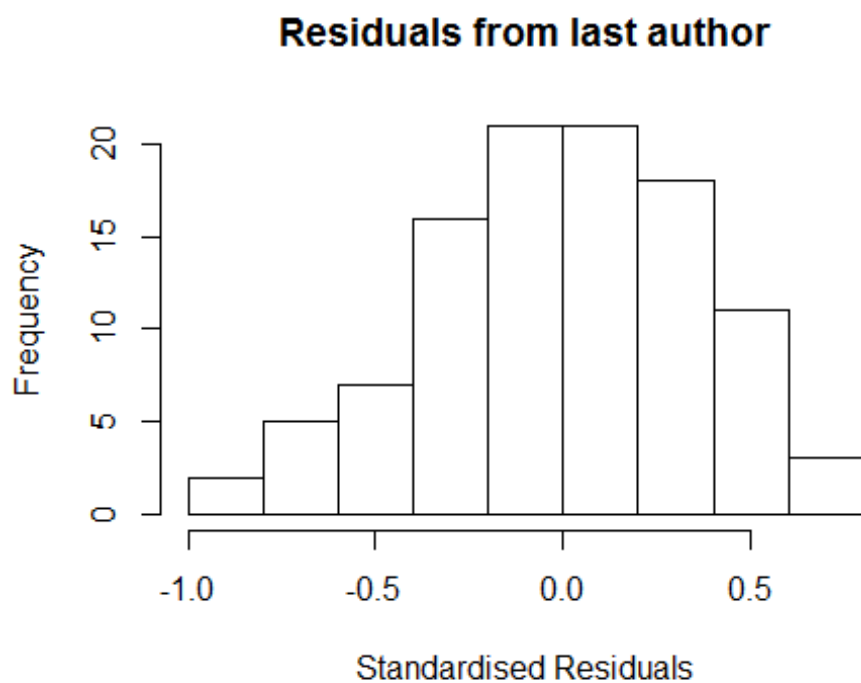
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2600"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 12 2 15 4 11 15 14 10 15 18 16 16 18 19
## 2011 2012
## 17 23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 8 0 8 1 4 10 10 3 10 10 8 9 10 11
## 2011 2012
## 13 14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 8 0 7 1 4 9 9 2 8 9 8 9 7 9
## 2011 2012
## 11 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.68179283050743"
## [1] "Male first author team size 2018 geometric mean: 1.16801609655919"

## 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 = 48, p-value = 0.06
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.68179283050743"
## [1] "Male last author team size 2018 geometric mean: 1.16801609655919"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

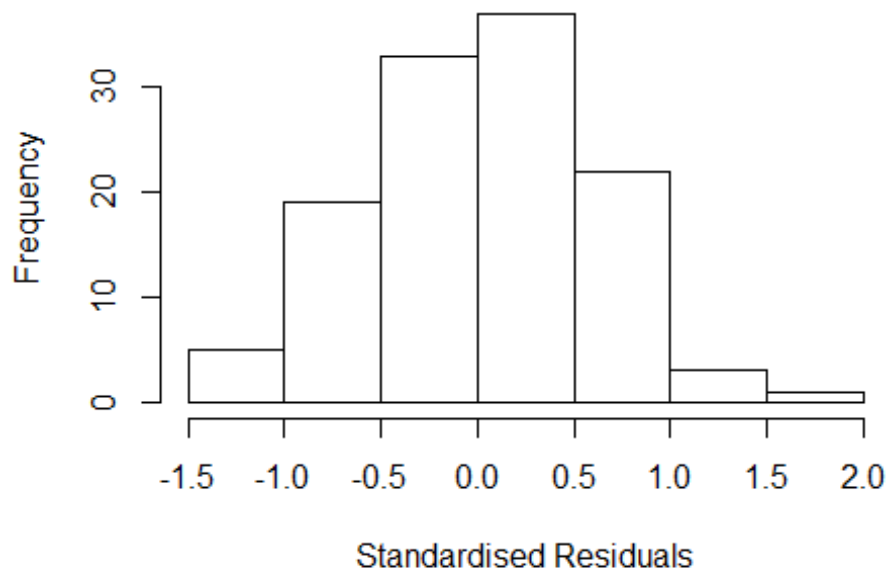
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 48, 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	1.675e+00	1	1.294
## LastAuthorFemale	-1.189e+15	1	NaN
## UniqueAuthors	-2.703e+16	4	NaN
## Year	-1.319e+16	15	NaN

Residuals from first and last author and team size



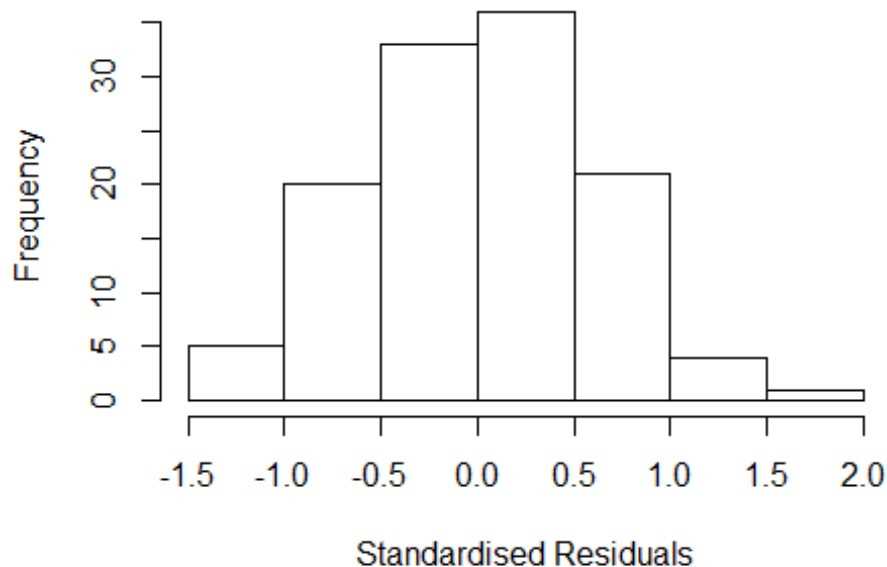
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4981 -0.3798 0.0276 0.4455 1.8834
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.95247 0.22684 4.20 5.9e-05 ***
## FirstAuthorFemale1 0.33897 0.26656 1.27 0.20651
## LastAuthorFemale1 0.04315 0.25050 0.17 0.86359
## UniqueAuthors2 0.17376 0.18509 0.94 0.35014
## UniqueAuthors3 -0.00592 0.20052 -0.03 0.97650
## UniqueAuthors4 -0.27230 0.38120 -0.71 0.47672
## UniqueAuthors5 1.00744 0.28551 3.53 0.00064 ***
## Year1997 -0.08987 0.28614 -0.31 0.75414
## Year1999 -0.23382 0.31791 -0.74 0.46380
## Year2000 0.28777 0.25323 1.14 0.25857
```

```

## Year2001      -0.22248    0.38093   -0.58  0.56053
## Year2002      0.20941    0.31797    0.66  0.51171
## Year2003      0.09080    0.25771    0.35  0.72535
## Year2004      0.13577    0.64465    0.21  0.83363
## Year2005      0.00518    0.30373    0.02  0.98643
## Year2006     -0.18875    0.28067   -0.67  0.50284
## Year2007     -0.61562    0.30842   -2.00  0.04870 *
## Year2008      0.14797    0.27614    0.54  0.59327
## Year2009      0.37189    0.40985    0.91  0.36643
## Year2010     -0.25307    0.26646   -0.95  0.34458
## Year2011      0.13742    0.31923    0.43  0.66780
## Year2012      0.44119    0.31220    1.41  0.16077
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.686
## Multiple R-squared:  0.201, Adjusted R-squared:  0.0296
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.431  0.907  0.959  0.928  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      8.33e-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 4.750 1 2.180
## LastAuthorFemale 3.978 1 1.994
## Year 2.466 15 1.031

```

Residuals from first and last author



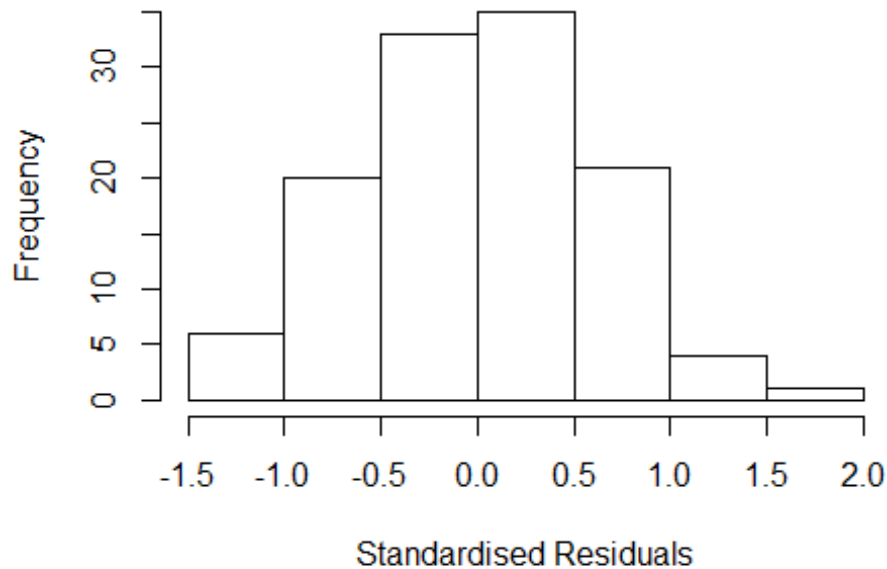
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4512 -0.4014 0.0151 0.4436 1.8555
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.98832 0.20202 4.89 3.7e-06 ***
## FirstAuthorFemale1 0.20197 0.24122 0.84 0.404
## LastAuthorFemale1 0.19758 0.23078 0.86 0.394
## Year1997 -0.12590 0.26708 -0.47 0.638
## Year1999 -0.26982 0.30158 -0.89 0.373
## Year2000 0.42568 0.20202 2.11 0.038 *
## Year2001 -0.20873 0.35533 -0.59 0.558
## Year2002 0.21528 0.31132 0.69 0.491
## Year2003 0.05680 0.23734 0.24 0.811
## Year2004 0.27368 0.62665 0.44 0.663
## Year2005 -0.00283 0.29683 -0.01 0.992
## Year2006 -0.20098 0.27020 -0.74 0.459
```

```

## Year2007          -0.62438    0.30146   -2.07    0.041 *
## Year2008          0.16867    0.27375    0.62    0.539
## Year2009          0.35221    0.38776    0.91    0.366
## Year2010         -0.20085    0.25993   -0.77    0.441
## Year2011          0.13271    0.30266    0.44    0.662
## Year2012          0.46286    0.30519    1.52    0.132
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.682
## Multiple R-squared:  0.174, Adjusted R-squared:  0.0366
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.440  0.898  0.955  0.926  0.990  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      8.33e-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 2.02 1          1.421
## Year              2.02 15          1.024

```

Residuals from first author



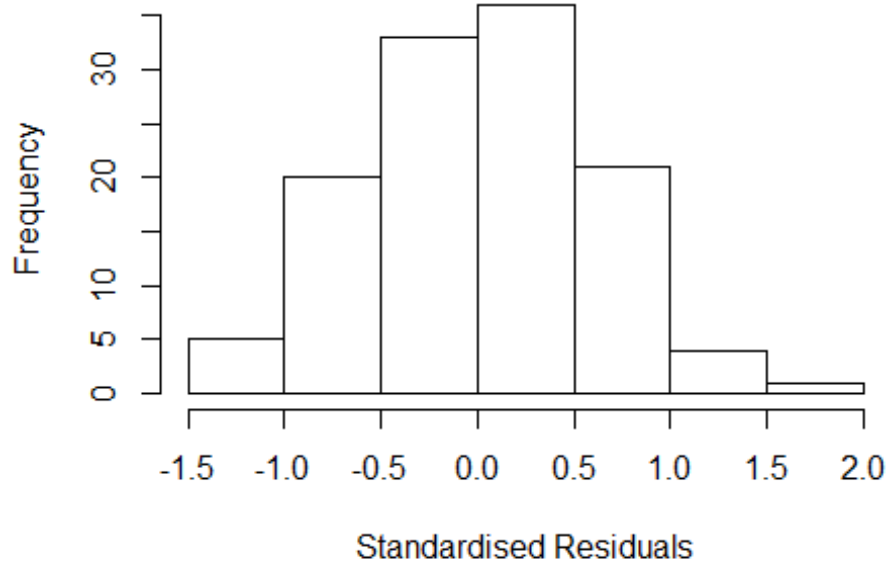
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4477 -0.3999 0.0324 0.4238 1.8556
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0086 0.2026 4.98 2.6e-06 ***
## FirstAuthorFemale1 0.3174 0.1688 1.88 0.063 .
## Year1997 -0.1462 0.2675 -0.55 0.586
## Year1999 -0.2901 0.3022 -0.96 0.339
## Year2000 0.4054 0.2026 2.00 0.048 *
## Year2001 -0.2290 0.3555 -0.64 0.521
## Year2002 0.1950 0.3115 0.63 0.533
## Year2003 0.0563 0.2328 0.24 0.809
## Year2004 0.2534 0.6269 0.40 0.687
## Year2005 -0.0231 0.2972 -0.08 0.938
## Year2006 -0.2078 0.2700 -0.77 0.443
## Year2007 -0.6378 0.3067 -2.08 0.040 *
```

```

## Year2008          0.1617      0.2725      0.59      0.554
## Year2009          0.3320      0.3879      0.86      0.394
## Year2010         -0.1749      0.2732     -0.64      0.523
## Year2011          0.1465      0.3122      0.47      0.640
## Year2012          0.4391      0.3081      1.43      0.157
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.682
## Multiple R-squared:  0.168, Adjusted R-squared:  0.039
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 111 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.439  0.895  0.959  0.927  0.990  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      8.33e-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 1.714 1          1.309
## Year            1.714 15          1.018

```


Residuals from last author



```
## [1] "List of 0 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
## -1.49073 -0.41209 -0.00847 0.42236 1.85795
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0100 0.2024 4.99 2.5e-06 ***
## LastAuthorFemale1 0.3119 0.1617 1.93 0.057 .
## Year1997 -0.1479 0.2672 -0.55 0.581
## Year1999 -0.2918 0.3020 -0.97 0.336
## Year2000 0.4040 0.2024 2.00 0.049 *
## Year2001 -0.2301 0.3562 -0.65 0.520
## Year2002 0.1942 0.3123 0.62 0.535
## Year2003 0.0250 0.2425 0.10 0.918
## Year2004 0.2520 0.6297 0.40 0.690
## Year2005 -0.0269 0.2957 -0.09 0.928
## Year2006 -0.2202 0.2771 -0.79 0.429
## Year2007 -0.5979 0.2933 -2.04 0.044 *
```

```

## Year2008          0.1520      0.2825      0.54      0.592
## Year2009          0.3313      0.3892      0.85      0.397
## Year2010         -0.2012      0.2596     -0.78      0.440
## Year2011          0.1409      0.2980      0.47      0.637
## Year2012          0.4807      0.2999      1.60      0.112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.675
## Multiple R-squared:  0.172, Adjusted R-squared:  0.0437
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.429  0.891  0.957   0.924  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      8.33e-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: 120"
## [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 1999 2000 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011 2012
##    3    3    2    1   10    3    2    2    3    8    7    5    7   18   12
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011 2012
##    2    1    1    1    5    1    1    0    3    5    2    4    4   15    7
##
## 1996 1997 1999 2000 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011 2012
##    2    1    0    0    5    1    1    0    3    5    2    3    3   11    6
## [1] "Heteroscedasticity checks, confirming that there are problems with

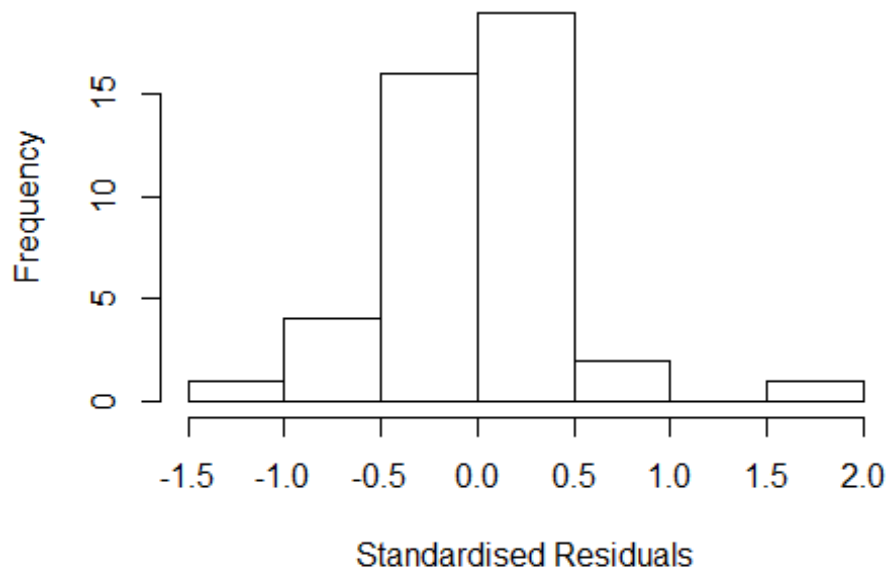
```

```

these"
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 2"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -9.015e+13  1      NaN
## LastAuthorFemale  -3.534e+00  1      NaN
## UniqueAuthors    -1.025e+18  4      NaN
## Year              -2.685e+05 11      NaN

```

Residuals from first and last author and team size



```

## [1] "List of 0 outliers with 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
## -1.15860 -0.21129  0.00289  0.23397  1.70826
##
## Coefficients:

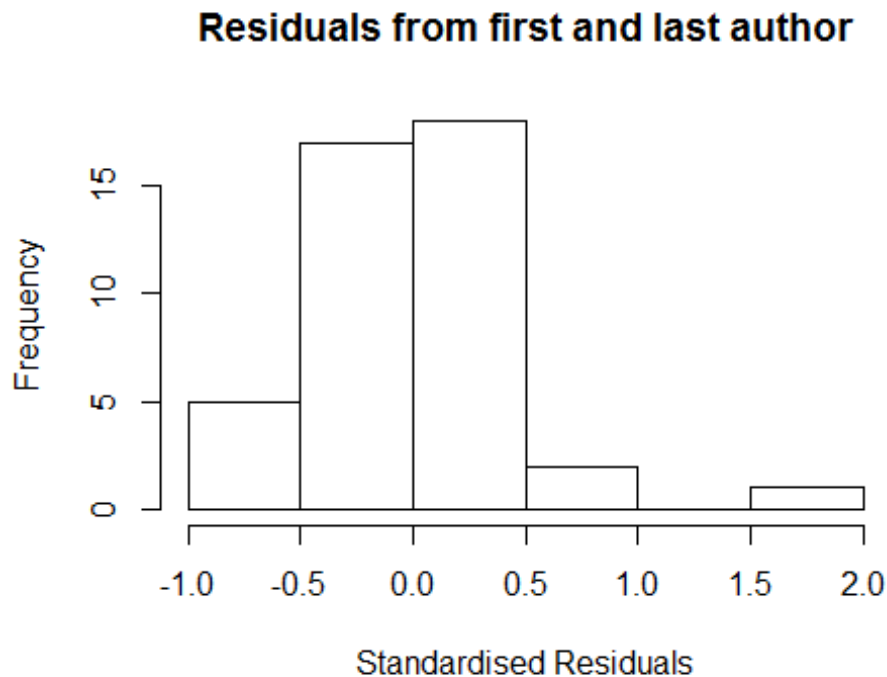
```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1563    0.2380   4.86 5.4e-05 ***
## FirstAuthorFemale1 0.3153    0.1482   2.13 0.0435 *
## LastAuthorFemale1 0.0432    0.4434   0.10 0.9232
## UniqueAuthors2    0.2457    0.2426   1.01 0.3209
## UniqueAuthors3   -0.1213    0.2571  -0.47 0.6412
## UniqueAuthors4   -0.2846    0.4140  -0.69 0.4981
## UniqueAuthors5   -0.3436    0.2471  -1.39 0.1766
## Year1997         -1.1563    0.2380  -4.86 5.4e-05 ***
## Year2001         -0.4382    0.3904  -1.12 0.2724
## Year2002         -0.5163    0.2380  -2.17 0.0398 *
## Year2003         -0.5353    0.2380  -2.25 0.0336 *
## Year2005         -0.6042    0.2145  -2.82 0.0093 **
## Year2007         -0.6018    0.3787  -1.59 0.1247
## Year2008         -0.7122    0.4571  -1.56 0.1318
## Year2009         -0.4361    0.2195  -1.99 0.0580 .
## Year2010         -0.5586    0.2731  -2.05 0.0515 .
## Year2011         -0.0380    0.2573  -0.15 0.8839
## Year2012        -0.8435    0.3263  -2.58 0.0160 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.409, Adjusted R-squared:  0.00757
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 33 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.124  0.904  0.965  0.903  0.981  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      2.33e-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 2.661 1          1.631

```

## LastAuthorFemale	3.464	1	1.861
## Year	3.651	11	1.061

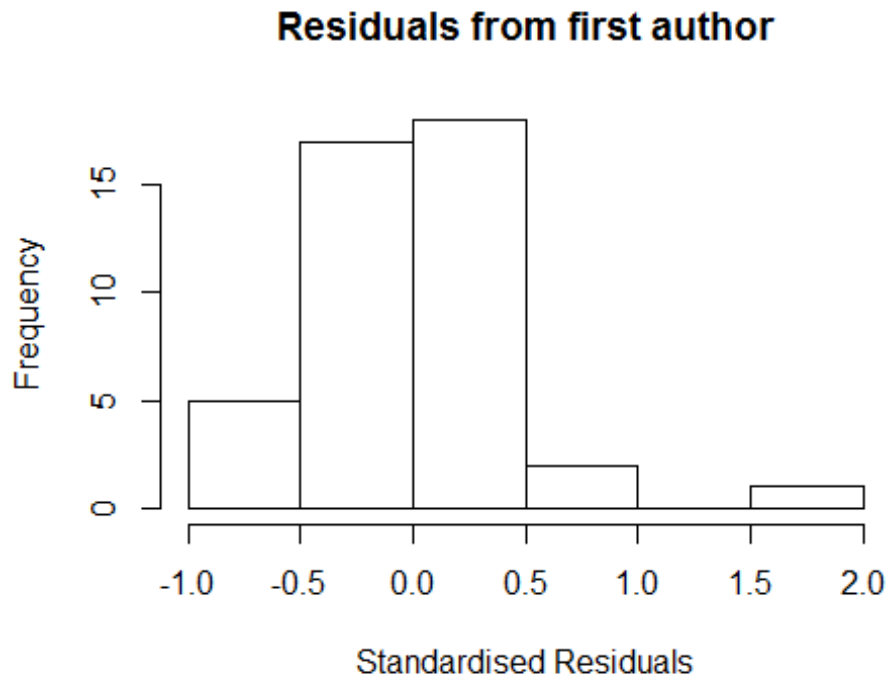


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.64e-01 -2.00e-01 -2.22e-16 3.00e-01 1.54e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2122 0.1705 7.11 8e-08 ***
## FirstAuthorFemale1 0.2684 0.1588 1.69 0.10177
## LastAuthorFemale1 -0.0217 0.2638 -0.08 0.93489
## Year1997 -1.2122 0.1705 -7.11 8e-08 ***
## Year2001 -0.3260 0.5589 -0.58 0.56421
## Year2002 -0.5722 0.1705 -3.36 0.00222 **
## Year2003 -0.5912 0.1705 -3.47 0.00166 **
## Year2005 -0.5788 0.1842 -3.14 0.00384 **
## Year2007 -0.4952 0.3601 -1.38 0.17954
```

```

## Year2008          -0.6128      0.7621   -0.80   0.42788
## Year2009          -0.3104      0.2030   -1.53   0.13703
## Year2010          -0.4248      0.3378   -1.26   0.21861
## Year2011          -0.1689      0.1627   -1.04   0.30778
## Year2012          -0.7302      0.1972   -3.70   0.00089 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.408
## Multiple R-squared:  0.356, Adjusted R-squared:  0.0671
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.125  0.899  0.961  0.896  0.986  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      2.33e-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 1.684 1          1.298
## Year              1.684 11          1.024

```



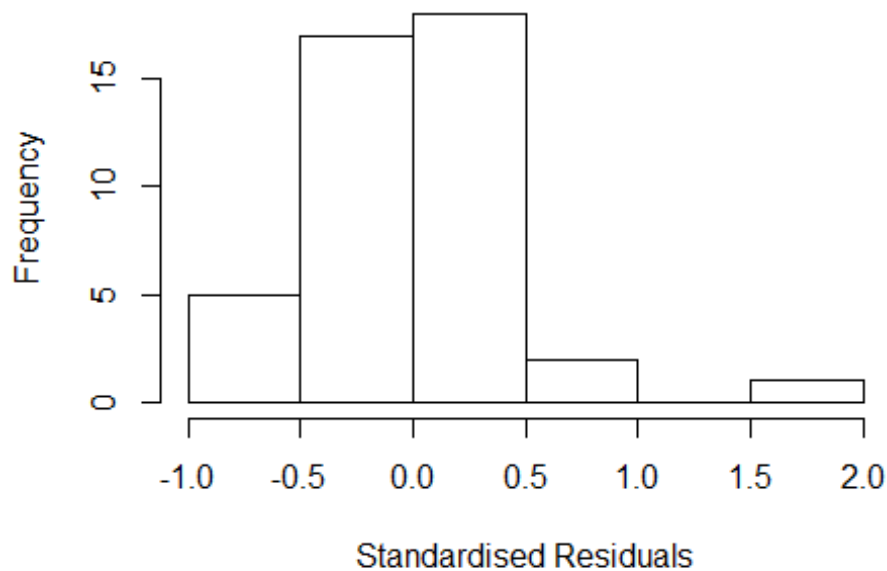
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.82e-01 -1.99e-01  1.11e-16  2.98e-01  1.55e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.206      0.144    8.35 2.6e-09 ***
## FirstAuthorFemale1  0.260      0.161    1.61 0.11712
## Year1997        -1.206      0.144   -8.35 2.6e-09 ***
## Year2001         -0.316      0.570   -0.55 0.58304
## Year2002         -0.566      0.144   -3.92 0.00048 ***
## Year2003         -0.585      0.144   -4.05 0.00033 ***
## Year2005         -0.572      0.160   -3.57 0.00123 **
## Year2007         -0.483      0.347   -1.39 0.17435
## Year2008         -0.617      0.758   -0.81 0.42192
## Year2009         -0.301      0.167   -1.80 0.08139 .
## Year2010         -0.410      0.326   -1.26 0.21762
## Year2011         -0.164      0.148   -1.11 0.27766
```

```

## Year2012          -0.736      0.200   -3.68  0.00091 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.398
## Multiple R-squared:  0.365, Adjusted R-squared:  0.111
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0957 0.8980 0.9590 0.8920 0.9840 0.9970
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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 4.067 1      2.017
## Year            4.067 11      1.066

```


Residuals from last author



```
## [1] "List of 0 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
## -9.28e-01 -2.49e-01 2.22e-16 3.19e-01 1.45e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.261 0.141 8.92 6.1e-10 ***
## LastAuthorFemale1 0.148 0.237 0.63 0.53562
## Year1997 -1.261 0.141 -8.92 6.1e-10 ***
## Year2001 -0.374 0.479 -0.78 0.44126
## Year2002 -0.621 0.141 -4.39 0.00013 ***
## Year2003 -0.640 0.141 -4.53 8.8e-05 ***
## Year2005 -0.628 0.157 -3.99 0.00040 ***
## Year2007 -0.481 0.320 -1.50 0.14331
## Year2008 -0.747 0.474 -1.58 0.12549
## Year2009 -0.273 0.224 -1.22 0.23337
## Year2010 -0.325 0.288 -1.13 0.26788
## Year2011 -0.169 0.155 -1.09 0.28512
```

```

## Year2012          -0.695      0.215   -3.23  0.00300 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared:  0.277, Adjusted R-squared:  -0.0116
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 36 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.336  0.895  0.953  0.913  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      2.33e-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: 43"
## [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
##    3    4    9    3    2    7    5    7    3    6    7    7    5   10    4
## 2011 2012
##    5   10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    5    0    0    4    3    3    2    3    5    5    4    9    4
## 2011 2012
##    2    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    5    0    0    4    3    3    1    3    4    4    3    8    4
## 2011 2012

```

```
##      2      5
## [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: 1.09050773266526"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -7.146e+15  1          NaN
## LastAuthorFemale -2.105e+15  1          NaN
## UniqueAuthors    -6.964e+29  2          NaN
## Year              6.948e+44 14         39.95

## [1] "List of 0 outliers with 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
## -9.57e-01 -2.35e-01 -4.44e-16  2.61e-01  1.01e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.18e+00   3.73e-08  3.17e+07 < 2e-16 ***
## FirstAuthorFemale1  3.93e-01   4.59e-01  8.60e-01  0.398
## LastAuthorFemale1  1.54e-01   5.15e-01  3.00e-01  0.767
## UniqueAuthors2     8.03e-02   1.87e-01  4.30e-01  0.671
## UniqueAuthors4    -6.53e-01   5.93e-01 -1.10e+00  0.278
## Year1997          -1.03e+00   1.41e-01 -7.36e+00 1.6e-08 ***
## Year1998          -2.29e-01   1.92e-01 -1.19e+00  0.242
## Year2001          -1.44e-01   1.98e-01 -7.20e-01  0.474
## Year2002          -1.15e-01   3.76e-01 -3.10e-01  0.762
## Year2003          -5.29e-02   1.51e-01 -3.50e-01  0.728
## Year2004          -5.65e-01   4.18e-08 -1.35e+07 < 2e-16 ***
## Year2005          -5.01e-01   1.96e-01 -2.55e+00  0.015 *
## Year2006          -1.04e-01   3.61e-01 -2.90e-01  0.775
## Year2007          -4.95e-01   2.28e-01 -2.17e+00  0.037 *
## Year2008          -2.38e-02   4.65e-01 -5.00e-02  0.960
## Year2009          -3.51e-01   1.41e-01 -2.49e+00  0.018 *
## Year2010          -1.70e-01   2.31e-01 -7.40e-01  0.466
## Year2011          -5.50e-01   1.02e-01 -5.40e+00 5.3e-06 ***
## Year2012          -2.24e-01   4.20e-01 -5.30e-01  0.597
## ---
```

```

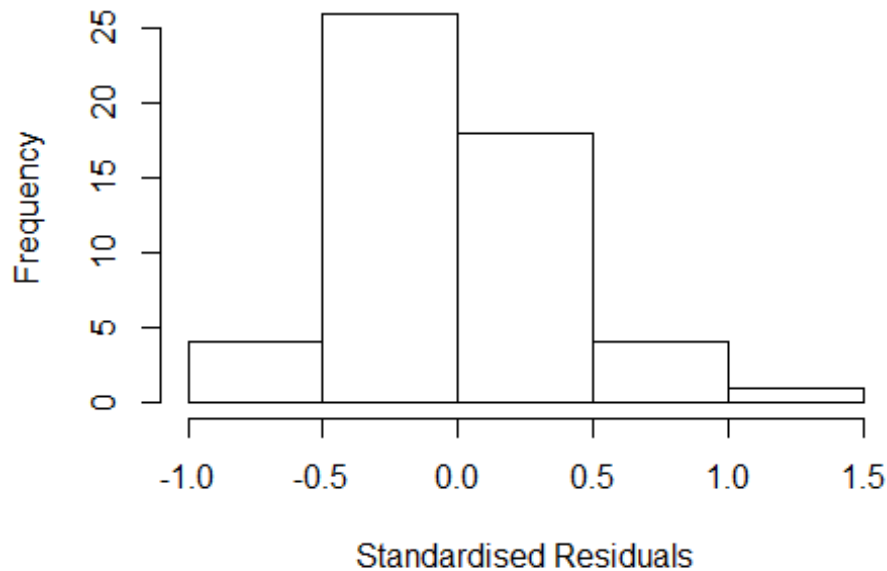
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.547
## Multiple R-squared:  0.228, Adjusted R-squared:  -0.181
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.715  0.936  0.973   0.944   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      1.89e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

```

Residuals from first and last author and team size



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## LastAuthorFemale  NaN 1             NaN
## Year              NaN 14            NaN

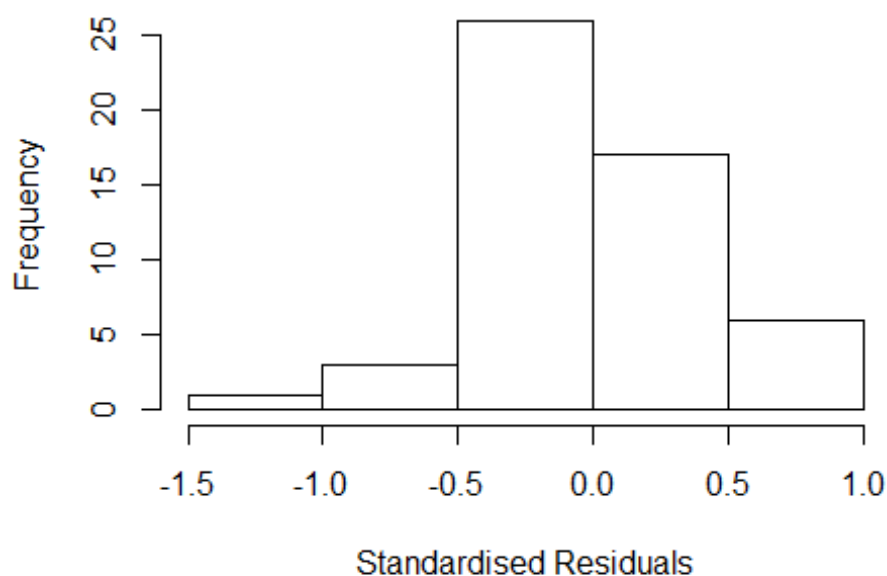
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0515 -0.2255 -0.0132  0.2567  0.9115
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1810     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1  0.1323     0.3366     0.39  0.697
## LastAuthorFemale1  0.3742     0.3624     1.03  0.309
## Year1997          -0.9945     0.1344    -7.40 9.8e-09 ***
## Year1998          -0.1797     0.1579    -1.14  0.262
## Year2001          -0.1031     0.1739    -0.59  0.557
## Year2002          -0.2610     0.3278    -0.80  0.431
```

```

## Year2003          -0.0525      0.1504    -0.35      0.729
## Year2004          -0.5650      0.0000    -Inf    < 2e-16 ***
## Year2005          -0.4478      0.1804    -2.48      0.018 *
## Year2006          -0.0636      0.3134    -0.20      0.840
## Year2007          -0.4357      0.1884    -2.31      0.027 *
## Year2008           0.0142      0.4870      0.03      0.977
## Year2009          -0.3496      0.1396    -2.50      0.017 *
## Year2010          -0.1700      0.2286    -0.74      0.462
## Year2011          -0.5505      0.1019    -5.40    4.3e-06 ***
## Year2012          -0.1295      0.3681    -0.35      0.727
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.206, Adjusted R-squared:  -0.147
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.728  0.942  0.977   0.951  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           1.89e-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"
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

```

Residuals from first and last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.381  1      1.839
## Year              3.381 14      1.044

## [1] "List of  0 outliers 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
## -1.0179 -0.2677 -0.0128  0.2565  0.9451
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.18e+00   0.00e+00      Inf < 2e-16 ***
## FirstAuthorFemale1  2.58e-01   2.39e-01   1.08e+00   0.286
## Year1997         -9.10e-01   1.42e-01  -6.40e+00  1.8e-07 ***
## Year1998         -1.79e-01   1.59e-01  -1.13e+00   0.265
## Year2001         -1.03e-01   1.74e-01  -5.90e-01   0.558
## Year2002         -3.08e-01   3.18e-01  -9.70e-01   0.339
## Year2003         -5.26e-02   1.51e-01  -3.50e-01   0.729
## Year2004         -5.65e-01   3.08e-08 -1.84e+07 < 2e-16 ***
```

```

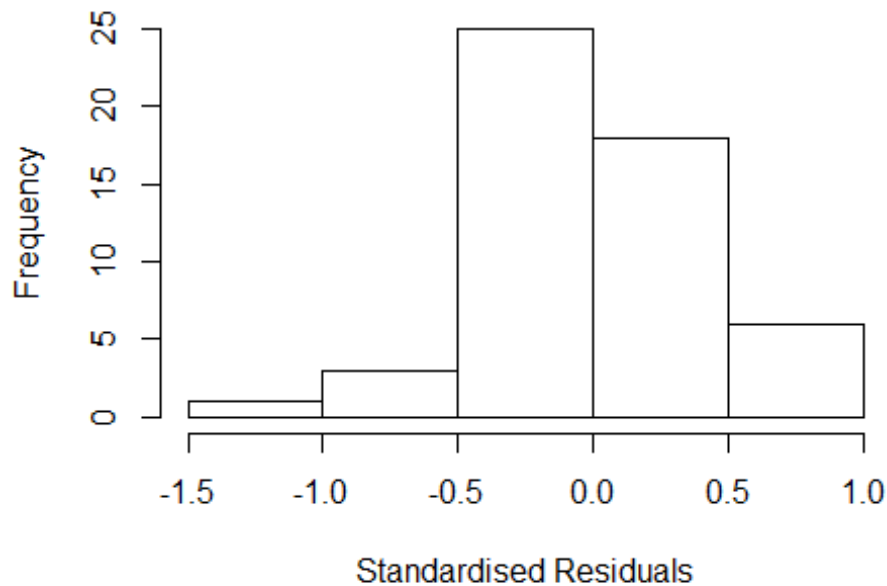
## Year2005          -4.48e-01    1.81e-01 -2.48e+00    0.018 *
## Year2006          -6.36e-02    3.16e-01 -2.00e-01    0.842
## Year2007          -4.36e-01    1.89e-01 -2.30e+00    0.027 *
## Year2008           1.38e-02    4.95e-01  3.00e-02    0.978
## Year2009          -3.50e-01    1.40e-01 -2.50e+00    0.017 *
## Year2010          -1.70e-01    2.29e-01 -7.40e-01    0.463
## Year2011          -5.50e-01    1.02e-01 -5.40e+00  4.1e-06 ***
## Year2012          -1.63e-01    3.47e-01 -4.70e-01    0.641
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.201, Adjusted R-squared:  -0.123
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~ = 1. The remaining 49 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.730  0.939  0.976  0.949  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      1.89e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

```


Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale NaN 1          NaN
## Year            NaN 14          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
## -1.0866 -0.2255 -0.0256  0.2947  0.8764
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.18e+00   2.35e-08  5.03e+07 < 2e-16 ***
## LastAuthorFemale1  5.06e-01   1.34e-01  3.77e+00  0.00057 ***
## Year1997        -9.95e-01   1.34e-01 -7.40e+00  8.3e-09 ***
## Year1998        -1.80e-01   1.58e-01 -1.14e+00  0.26233
## Year2001        -1.03e-01   1.74e-01 -5.90e-01  0.55670
## Year2002        -2.12e-01   2.45e-01 -8.70e-01  0.39159
## Year2003        -5.25e-02   1.50e-01 -3.50e-01  0.72906
## Year2004        -5.65e-01   0.00e+00  -Inf    < 2e-16 ***
```

```

## Year2005          -4.48e-01    1.80e-01 -2.48e+00  0.01774 *
## Year2006          -6.36e-02    3.14e-01 -2.00e-01  0.84042
## Year2007          -4.36e-01    1.88e-01 -2.31e+00  0.02642 *
## Year2008           1.42e-02    4.87e-01  3.00e-02  0.97698
## Year2009          -3.50e-01    1.40e-01 -2.50e+00  0.01680 *
## Year2010          -1.70e-01    2.29e-01 -7.40e-01  0.46191
## Year2011          -5.50e-01    1.02e-01 -5.40e+00  4.0e-06 ***
## Year2012          -9.44e-02    3.23e-01 -2.90e-01  0.77199
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.205, Adjusted R-squared:  -0.118
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~ = 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.710  0.942  0.976  0.951  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      1.89e-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: 53"
## [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
##    1    3    2    5    6    4    4    1    3    4    7    4    3    5    3
## 2011 2012
##    4    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

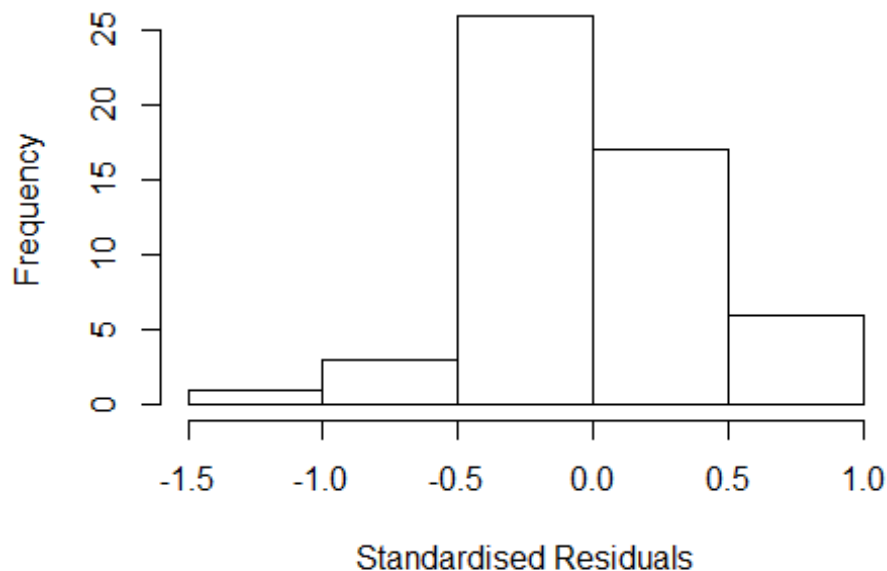
```

```
##      0      1      0      2      3      1      3      0      1      3      5      2      2      3      2
## 2011 2012
##      1      2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      1      0      2      2      1      2      0      0      3      5      2      2      2      2
## 2011 2012
##      0      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: 1"
## [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"

## 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
```

Residuals from last author

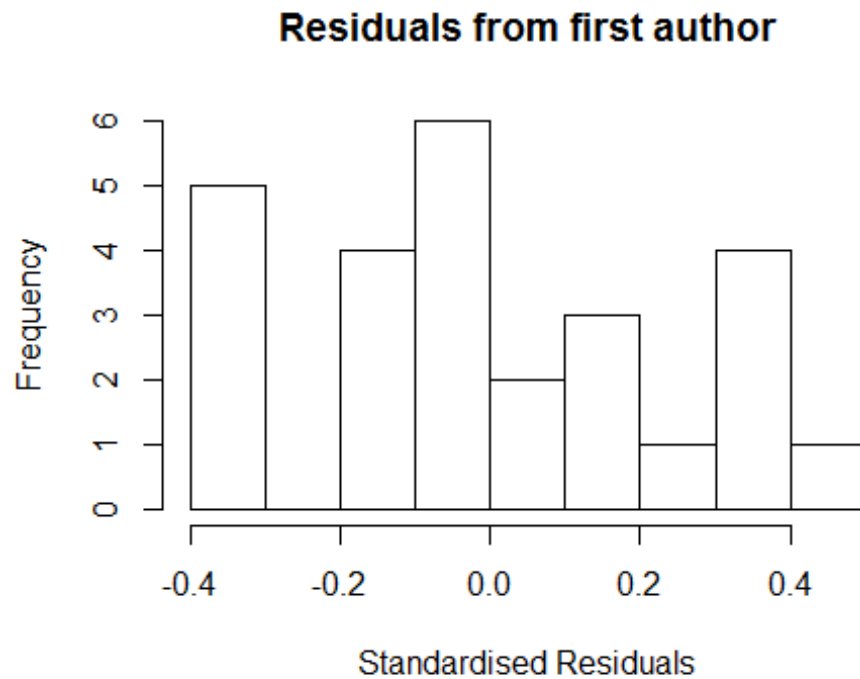


```
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1          NaN
## Year              NaN 11         NaN
```

```
## [1] "Regression 4: Last author gender, Year as factors"

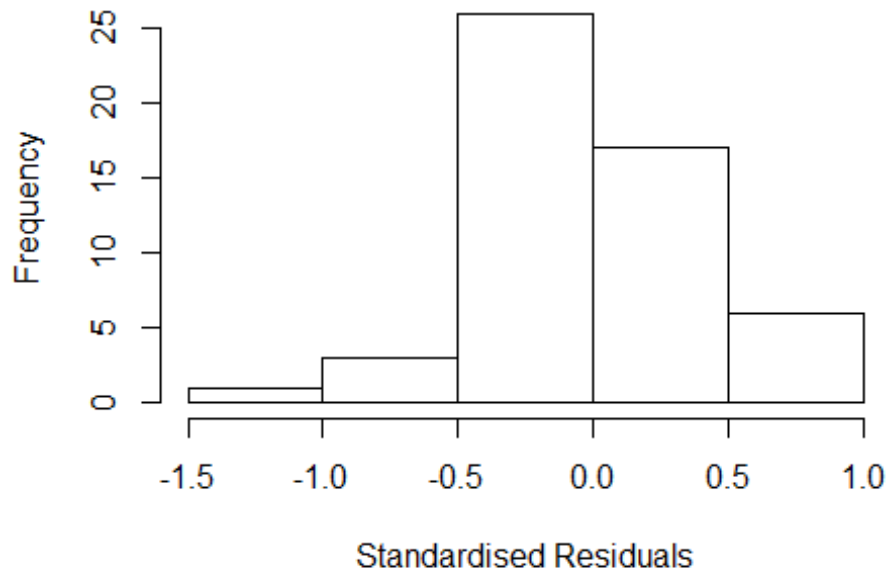
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```



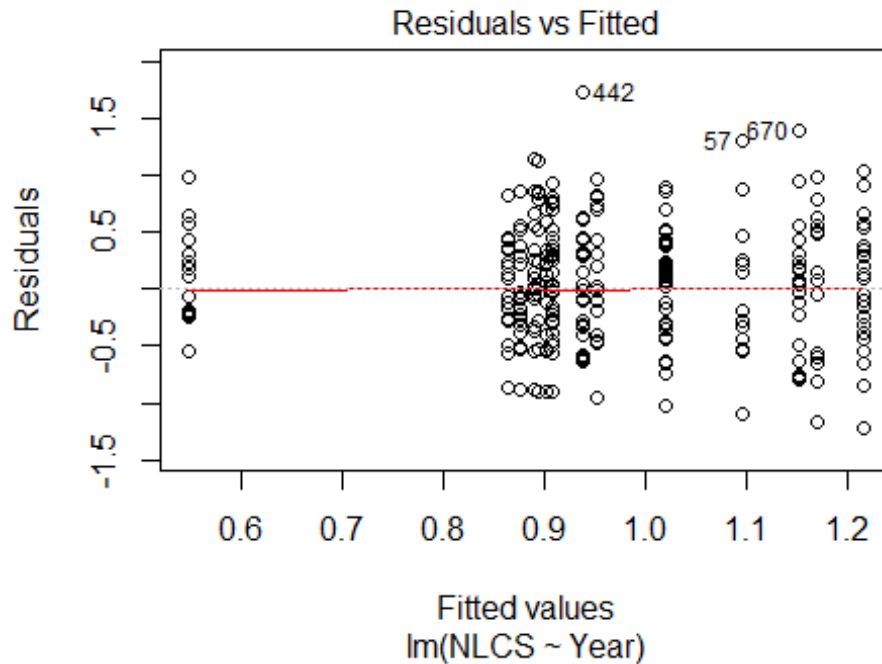
```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale NaN 1          NaN
## Year            NaN 11          NaN
```

Residuals from last author



```
## [1] "Sample size for the above analysis: 26"
## [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 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 27 32 39 2 55 36 30 32 27 39 42 31 41 42 49
## 2012
## 50
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 15 14 16 0 21 19 13 19 21 29 21 15 23 22 30
## 2012
## 30
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 14 13 14 0 17 18 13 16 21 27 17 14 18 20 25
## 2012
## 23
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
```

```
##
## data: NLCS by Year
## Bartlett's K-squared = 12, df = 14, p-value = 0.6
```



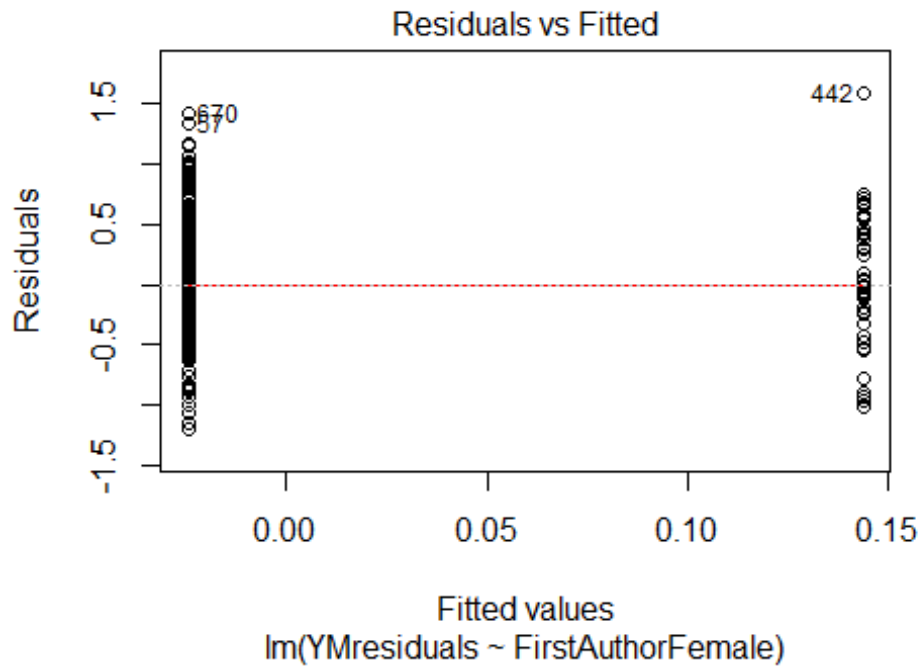
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.2, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male first author team size 2018 geometric mean: 1.93078203668652"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

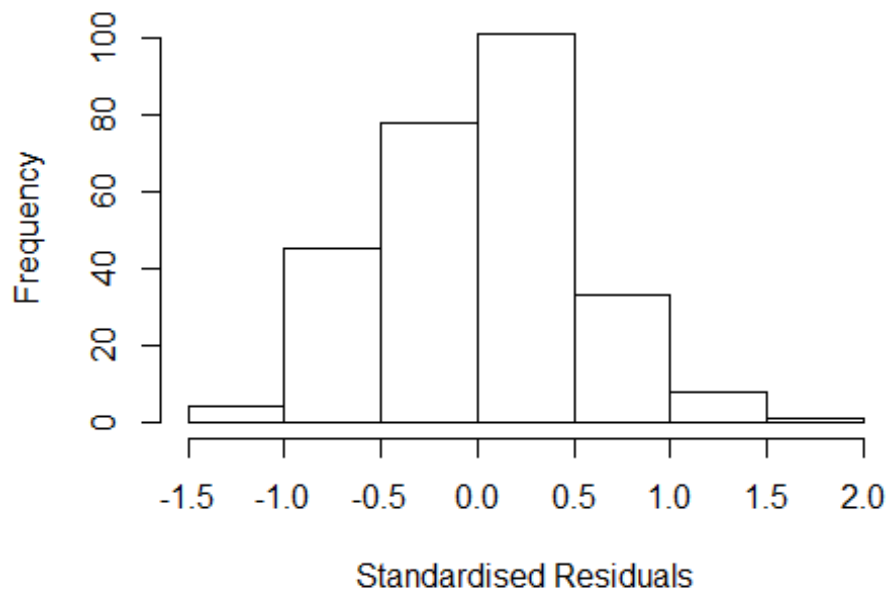
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 1.78067980738487"

## 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 = 15, 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 1.562 1          1.250
## LastAuthorFemale  1.463 1          1.210
## UniqueAuthors    3.271 4          1.160
## Year              3.598 14         1.047
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.1036 -0.3784 0.0286 0.3402 1.9262
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8994 0.2148 4.19 3.9e-05 ***
## FirstAuthorFemale1 0.1156 0.1115 1.04 0.3007
## LastAuthorFemale1 -0.2109 0.1127 -1.87 0.0625 .
## UniqueAuthors2 0.1208 0.0800 1.51 0.1321
## UniqueAuthors3 0.3815 0.1181 3.23 0.0014 **
## UniqueAuthors4 0.2667 0.2126 1.25 0.2110
## UniqueAuthors5 0.4067 0.1623 2.51 0.0129 *
## Year1997 0.0804 0.2694 0.30 0.7657
## Year1998 0.1598 0.3106 0.51 0.6073
## Year2001 -0.3999 0.2320 -1.72 0.0860 .
```

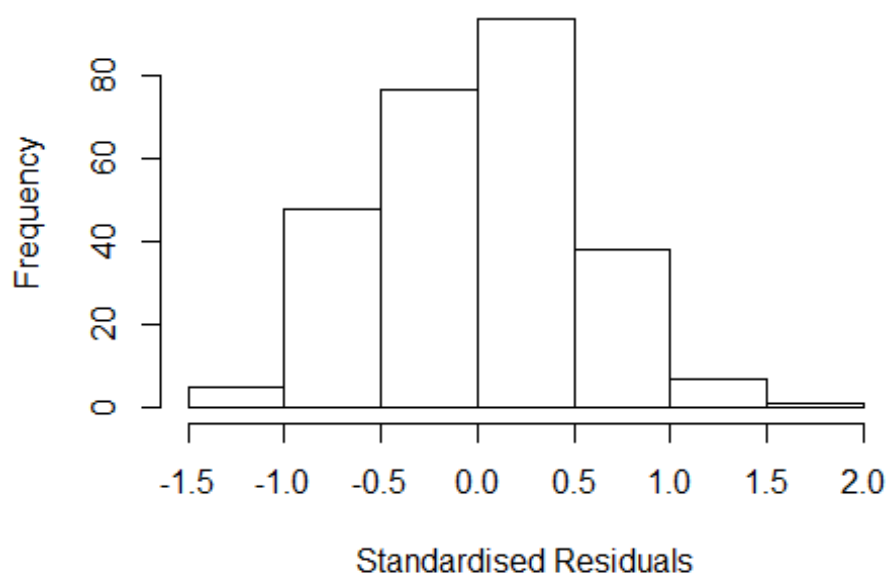


```

## Year2002          -0.1460      0.2517   -0.58   0.5624
## Year2003           0.0211      0.2635    0.08   0.9364
## Year2004          -0.0815      0.2849   -0.29   0.7751
## Year2005          -0.1390      0.2303   -0.60   0.5466
## Year2006          -0.0633      0.2258   -0.28   0.7793
## Year2007          -0.0946      0.2312   -0.41   0.6828
## Year2008          -0.0973      0.2457   -0.40   0.6924
## Year2009           0.1880      0.2580    0.73   0.4668
## Year2010           0.1051      0.2463    0.43   0.6699
## Year2011          -0.0820      0.2215   -0.37   0.7116
## Year2012          -0.0338      0.2326   -0.15   0.8847
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.131, Adjusted R-squared:  0.0617
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 246 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.159  0.870  0.945  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      3.70e-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 1.350 1      1.162
## LastAuthorFemale  1.182 1      1.087
## Year              1.596 14      1.017

```

Residuals from first and last author



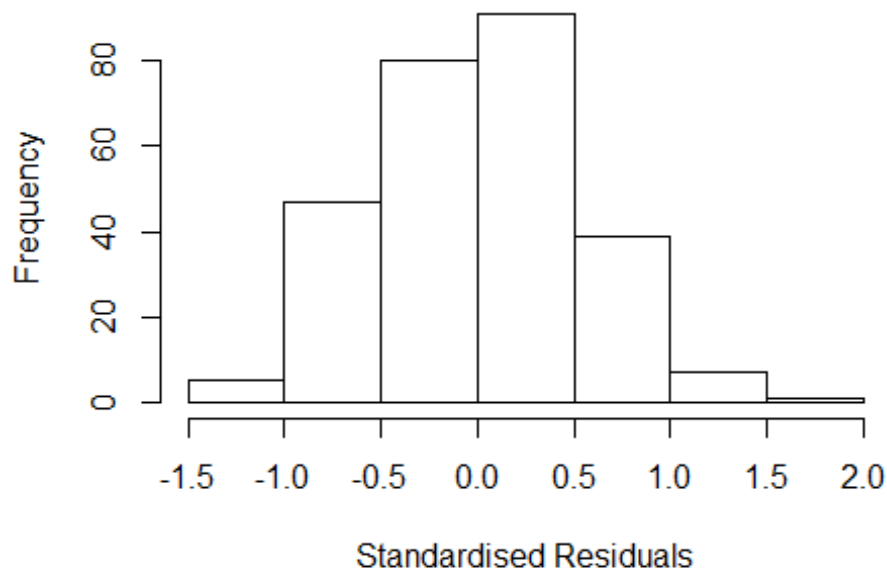
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2279 -0.3908 0.0273 0.3645 1.7264
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.96210 0.22489 4.28 2.7e-05 ***
## FirstAuthorFemale1 0.18964 0.10253 1.85 0.066 .
## LastAuthorFemale1 -0.15801 0.11235 -1.41 0.161
## Year1997 0.06290 0.28999 0.22 0.828
## Year1998 0.17291 0.32814 0.53 0.599
## Year2001 -0.37565 0.24428 -1.54 0.125
## Year2002 -0.10855 0.26439 -0.41 0.682
## Year2003 0.08806 0.27916 0.32 0.753
## Year2004 -0.01761 0.30353 -0.06 0.954
## Year2005 -0.09688 0.24367 -0.40 0.691
## Year2006 -0.05319 0.23776 -0.22 0.823
## Year2007 -0.12281 0.24841 -0.49 0.621
```

```

## Year2008      -0.07838    0.27151   -0.29    0.773
## Year2009      0.26576    0.27131    0.98    0.328
## Year2010      0.12740    0.25914    0.49    0.623
## Year2011      0.00433    0.23642    0.02    0.985
## Year2012     -0.03214    0.25053   -0.13    0.898
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.554
## Multiple R-squared:  0.0808, Adjusted R-squared:  0.0226
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 248 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.311  0.874  0.952  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      3.70e-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 1.385 1          1.177
## Year              1.385 14          1.012

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2230 -0.3854 0.0302 0.3741 1.6262
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.94850 0.22541 4.21 3.6e-05 ***
## FirstAuthorFemale1 0.15425 0.10962 1.41 0.16
## Year1997 0.07995 0.28926 0.28 0.78
## Year1998 0.18501 0.33419 0.55 0.58
## Year2001 -0.37262 0.24600 -1.51 0.13
## Year2002 -0.12202 0.26505 -0.46 0.65
## Year2003 0.09153 0.28072 0.33 0.74
## Year2004 -0.00151 0.30455 0.00 1.00
## Year2005 -0.09614 0.24645 -0.39 0.70
## Year2006 -0.06193 0.24003 -0.26 0.80
## Year2007 -0.13325 0.25239 -0.53 0.60
## Year2008 -0.06226 0.27282 -0.23 0.82
```

```

## Year2009          0.27451    0.27048    1.01    0.31
## Year2010          0.13950    0.25893    0.54    0.59
## Year2011          0.00754    0.23869    0.03    0.97
## Year2012         -0.03756    0.25287   -0.15    0.88
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.553
## Multiple R-squared:  0.0757, Adjusted R-squared:  0.0211
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 246 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.367  0.865  0.947  0.909  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      3.70e-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 1.187 1          1.089
## Year            1.187 14          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
## -1.2499 -0.3933  0.0291  0.3627  1.8483
##
## Coefficients:

```

```

##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9771    0.2241   4.36 1.9e-05 ***
## LastAuthorFemale1 -0.1027    0.1136  -0.90   0.37
## Year1997          0.0681    0.2887   0.24   0.81
## Year1998          0.1859    0.3220   0.58   0.56
## Year2001         -0.3943    0.2436  -1.62   0.11
## Year2002         -0.0965    0.2647  -0.36   0.72
## Year2003          0.0924    0.2857   0.32   0.75
## Year2004         -0.0206    0.3073  -0.07   0.95
## Year2005         -0.0991    0.2443  -0.41   0.69
## Year2006         -0.0557    0.2378  -0.23   0.81
## Year2007         -0.1129    0.2490  -0.45   0.65
## Year2008         -0.0815    0.2744  -0.30   0.77
## Year2009          0.2728    0.2733   1.00   0.32
## Year2010          0.1364    0.2566   0.53   0.60
## Year2011          0.0472    0.2367   0.20   0.84
## Year2012         -0.0368    0.2480  -0.15   0.88
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.555
## Multiple R-squared:  0.0705, Adjusted R-squared:  0.0156
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 253 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.245  0.875  0.947  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      3.70e-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: 270"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2605"
## [1] "#####"

```

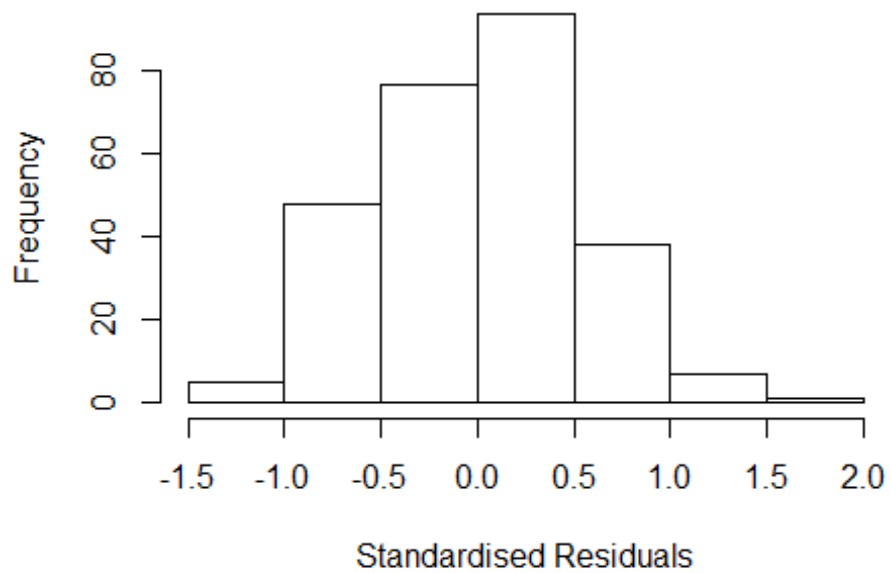
```

## [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
##    9   10    7    7   12   14    5   10   16   10    8    9    7    8   10
## 2011 2012
##    7   11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    2    2    1    6    3    5    9    8    5    5    5    4    4
## 2011 2012
##    5    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    2    2    1    5    3    5    6    7    4    4    5    2    4
## 2011 2012
##    5    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 1"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 1.41421356237309"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate

```

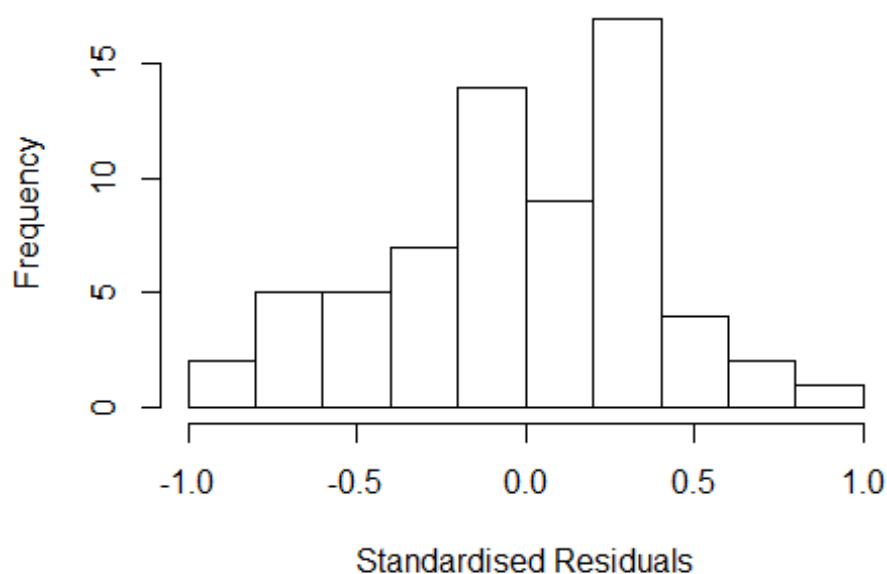
Residuals from last author



```
## [1] "Regression 2: First author gender, Last author gender, Year as factors"
```

##		GVIF	Df	$GVIF^{1/(2 \cdot Df)}$
##	FirstAuthorFemale	2.701	1	1.643
##	LastAuthorFemale	6.194	1	2.489
##	Year	11.311	16	1.079

Residuals from first and last author



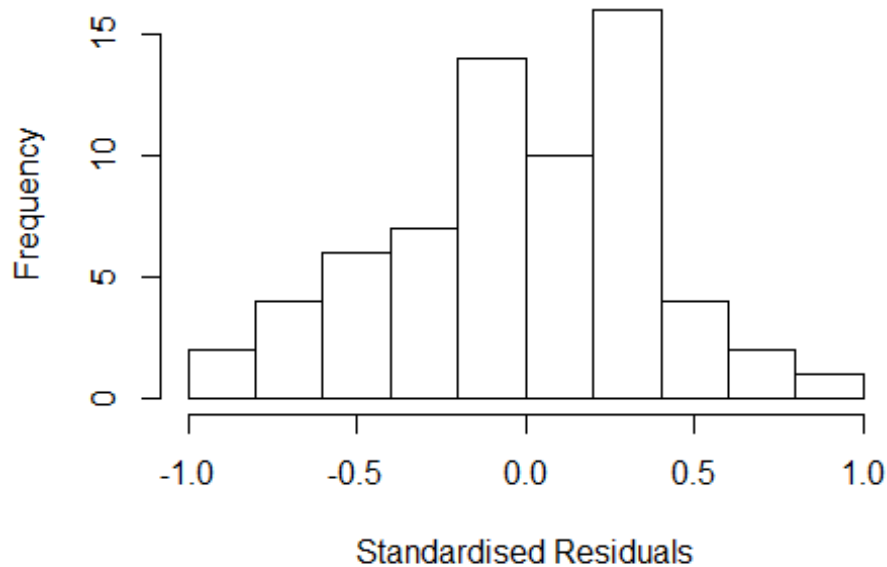
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8881 -0.2149 0.0044 0.2807 0.8151
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7545 0.3448 2.19 0.034 *
## FirstAuthorFemale1 0.2524 0.2102 1.20 0.236
## LastAuthorFemale1 0.0746 0.2667 0.28 0.781
## Year1997 0.1235 0.3873 0.32 0.751
## Year1998 -0.1295 0.7838 -0.17 0.869
## Year1999 -0.2130 0.3787 -0.56 0.576
## Year2000 -0.2915 0.3448 -0.85 0.402
## Year2001 -0.2349 0.3592 -0.65 0.516
## Year2002 -0.1159 0.3832 -0.30 0.764
## Year2003 0.3898 0.3664 1.06 0.293
## Year2004 0.2358 0.3835 0.61 0.542
## Year2005 0.0317 0.3699 0.09 0.932
```

```

## Year2006          -0.0613      0.3789   -0.16    0.872
## Year2007           0.1335      0.4203    0.32    0.752
## Year2008           0.3100      0.4817    0.64    0.523
## Year2009          -0.1300      0.3975   -0.33    0.745
## Year2010           0.5083      0.4278    1.19    0.241
## Year2011           0.2359      0.4135    0.57    0.571
## Year2012           0.0636      0.4182    0.15    0.880
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.455
## Multiple R-squared:  0.236, Adjusted R-squared:  -0.0562
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 58 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.682  0.880  0.966  0.926  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      1.52e-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 2.833 1      1.683
## Year              2.833 16      1.033

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.8902 -0.2148 0.0015 0.2754 0.8877
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7563 0.3466 2.18 0.034 *
## FirstAuthorFemale1 0.2659 0.2041 1.30 0.199
## Year1997 0.1170 0.3882 0.30 0.764
## Year1998 -0.1313 0.8060 -0.16 0.871
## Year1999 -0.2148 0.3803 -0.56 0.575
## Year2000 -0.2933 0.3466 -0.85 0.402
## Year2001 -0.2367 0.3609 -0.66 0.515
## Year2002 -0.0972 0.3704 -0.26 0.794
## Year2003 0.3857 0.3673 1.05 0.299
## Year2004 0.2474 0.3842 0.64 0.523
## Year2005 0.0357 0.3717 0.10 0.924
## Year2006 -0.0289 0.3552 -0.08 0.936
```

```

## Year2007          0.1339      0.4211      0.32      0.752
## Year2008          0.3067      0.4905      0.63      0.535
## Year2009         -0.1318      0.3992     -0.33      0.743
## Year2010          0.5075      0.4294      1.18      0.243
## Year2011          0.2360      0.4282      0.55      0.584
## Year2012          0.0626      0.4199      0.15      0.882
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.449
## Multiple R-squared:  0.235, Adjusted R-squared:  -0.0353
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.674  0.877  0.967   0.926  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      1.52e-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 3.093 1      1.759
## Year              3.093 16      1.036
##
## [1] "List of 0 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
## -0.891086 -0.215602  0.000658  0.270493  0.904229

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7571    0.3473    2.18  0.034 *
## LastAuthorFemale1 0.1331    0.2341    0.57  0.572
## Year1997        0.2083    0.3679    0.57  0.574
## Year1998       -0.1321    0.8005   -0.16  0.870
## Year1999       -0.2156    0.3811   -0.57  0.574
## Year2000       -0.2941    0.3473   -0.85  0.401
## Year2001       -0.2375    0.3617   -0.66  0.515
## Year2002       -0.0485    0.3725   -0.13  0.897
## Year2003        0.4314    0.3826    1.13  0.265
## Year2004        0.2231    0.3837    0.58  0.564
## Year2005        0.0496    0.3697    0.13  0.894
## Year2006       -0.0285    0.3708   -0.08  0.939
## Year2007        0.1340    0.4214    0.32  0.752
## Year2008        0.3307    0.5574    0.59  0.556
## Year2009       -0.1326    0.3999   -0.33  0.742
## Year2010        0.5071    0.4301    1.18  0.244
## Year2011        0.2278    0.4105    0.55  0.581
## Year2012        0.0622    0.4206    0.15  0.883
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.215, Adjusted R-squared:  -0.0626
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 57 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.661  0.882  0.965  0.920  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      1.52e-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: 66"
## [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]"
##
## 1997 1998 1999 2000 2002 2003 2004 2006 2008 2009 2010 2011 2012
##    1    1    1    2    3    4    2    2    2    1    2    1    7
##
## 1997 1998 1999 2000 2002 2003 2004 2006 2008 2009 2010 2011 2012
##    1    0    0    1    1    0    1    1    0    0    0    0    5
##
## 1997 1998 1999 2000 2002 2003 2004 2006 2008 2009 2010 2011 2012
##    1    0    0    0    1    0    1    1    0    0    0    0    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 2.6207413942089"

## 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 = 2, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 2.6207413942089"

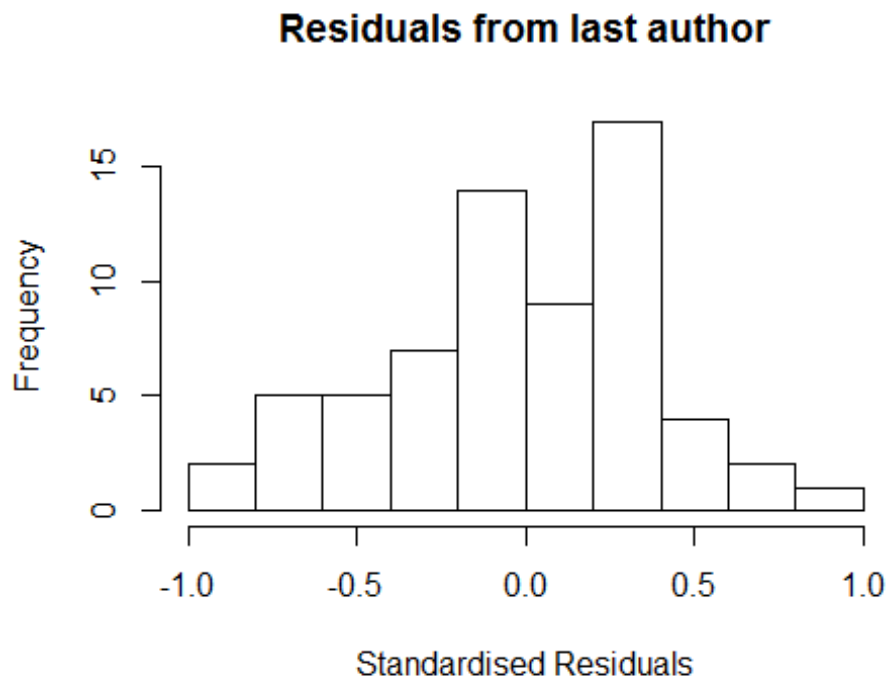
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2, 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"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
## [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

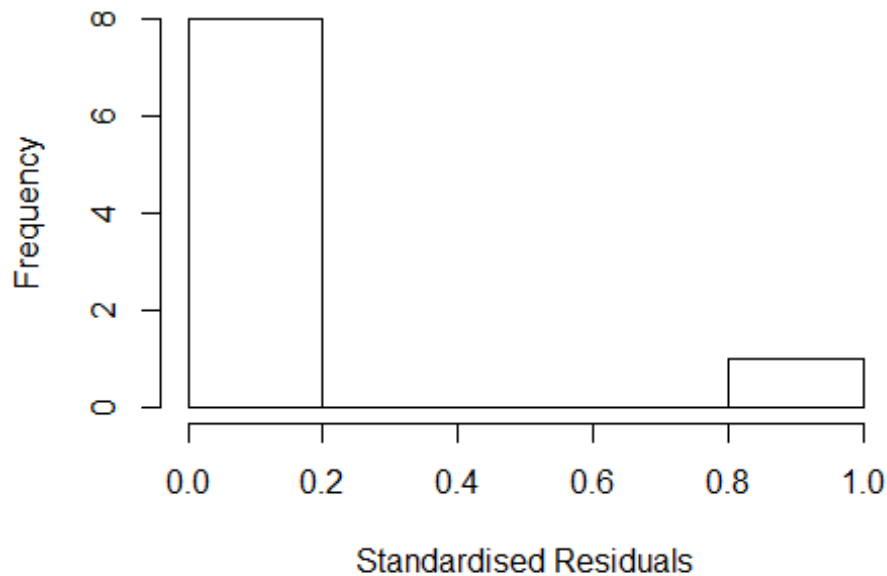
```

```
## 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	4	NaN

Residuals from first author

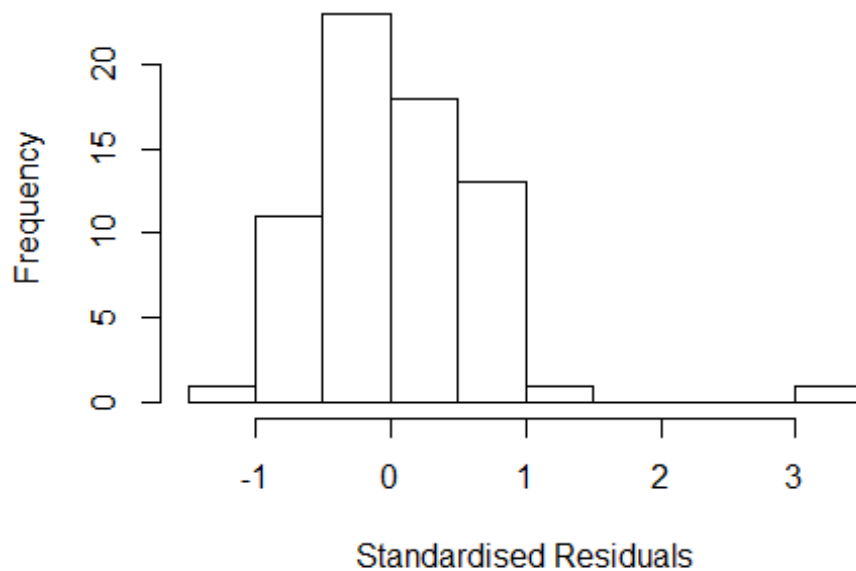


```
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 9"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2607"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    7    9    9    6    8    8    5    8    8    5    5    9    5
## 2011 2012
##    2    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    3    7    6    3    5    2    7    3    5    5    3    5    7    4
## 2011 2012
##    1    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    3    7    6    3    4    2    6    2    5    5    2    5    7    3
## 2011 2012
##    1    3
## [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: 1.31950791077289"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale  5.861  1          2.421
## LastAuthorFemale  3.600  1          1.897
## UniqueAuthors    8.347  2          1.700
## Year              35.491 16          1.118
```

Residuals from first and last author and team size



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 9 0030305457 4.238 1996    1804      3      3.051
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
## -1.214564 -0.374232 -0.000833  0.428853  3.051461
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9030    0.2757   3.28  0.0020 **
```

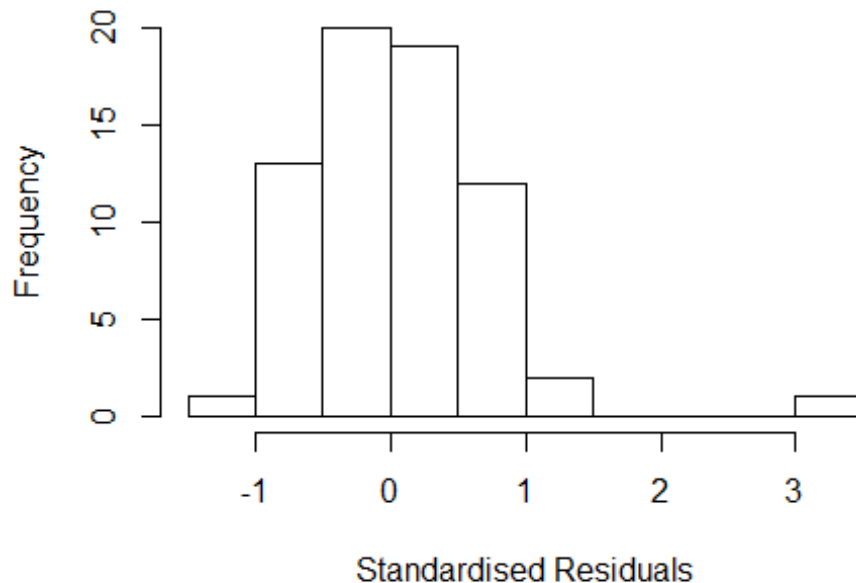
```

## FirstAuthorFemale1 -0.1090      0.2277    -0.48    0.6344
## LastAuthorFemale1 -0.1224      0.2208    -0.55    0.5820
## UniqueAuthors2     0.2836      0.2251     1.26    0.2140
## UniqueAuthors3     1.1255      0.3354     3.36    0.0016 **
## Year1997           -0.1643      0.3190    -0.52    0.6089
## Year1998           -0.4196      0.3680    -1.14    0.2599
## Year1999           -0.0216      0.3761    -0.06    0.9545
## Year2000           -0.2076      0.5683    -0.37    0.7166
## Year2001            0.3116      0.6240     0.50    0.6198
## Year2002           -0.2810      0.4915    -0.57    0.5702
## Year2003           -0.0351      0.4697    -0.07    0.9407
## Year2004            0.3227      0.4010     0.80    0.4250
## Year2005           -0.1199      0.3985    -0.30    0.7648
## Year2006            0.2566      0.3821     0.67    0.5052
## Year2007           -0.0520      0.4676    -0.11    0.9120
## Year2008           -0.1383      0.4438    -0.31    0.7567
## Year2009           -0.3681      0.3582    -1.03    0.3094
## Year2010           -0.4973      0.2758    -1.80    0.0778 .
## Year2011           -0.5190      0.2757    -1.88    0.0660 .
## Year2012           -0.5382      0.3569    -1.51    0.1382
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.717
## Multiple R-squared:  0.278, Adjusted R-squared:  -0.0297
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0303 0.9280 0.9620 0.9340 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      1.47e-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	16.43	1	4.054
## LastAuthorFemale	11.52	1	3.394
## Year	85.08	16	1.149

Residuals from first and last author



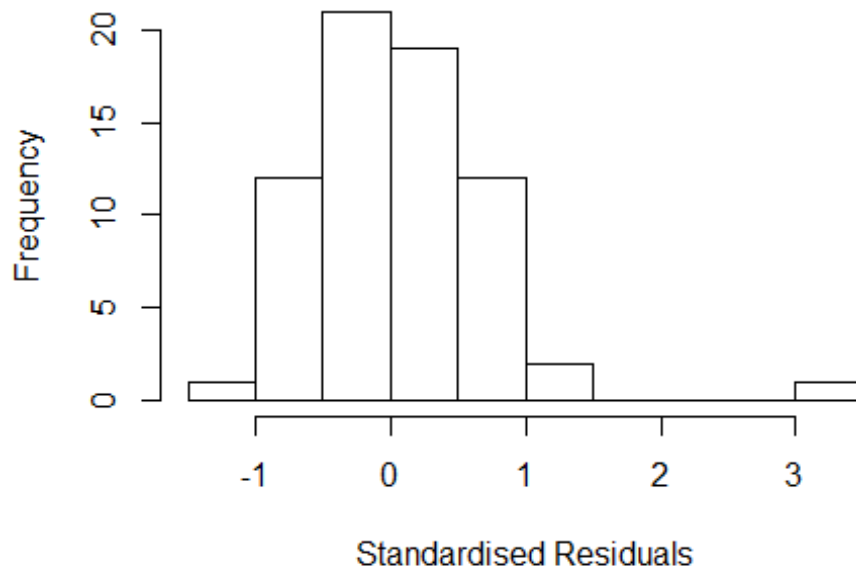
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 9 0030305457 4.238 1996    1804      3      3.362
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37503 -0.40767  0.00817  0.41346  3.36208
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.87592    0.22828   3.84 0.00036 ***
## FirstAuthorFemale1 0.11681    0.41982   0.28 0.78201
## LastAuthorFemale1 -0.38171    0.26634  -1.43 0.15816
## Year1997        -0.10682    0.34607  -0.31 0.75889
## Year1998        -0.10445    0.36880  -0.28 0.77821
## Year1999         0.05794    0.30887   0.19 0.85198
## Year2000         0.10570    0.47618   0.22 0.82526
## Year2001         0.49911    0.52431   0.95 0.34580
```

```

## Year2002          0.02958      0.33649      0.09  0.93031
## Year2003         -0.00955      0.41109     -0.02  0.98157
## Year2004          0.49158      0.25221      1.95  0.05702 .
## Year2005          0.03850      0.44748      0.09  0.93179
## Year2006          0.35580      0.33258      1.07  0.28996
## Year2007         -0.02492      0.43897     -0.06  0.95496
## Year2008          0.00362      0.38706      0.01  0.99257
## Year2009         -0.14895      0.33810     -0.44  0.66147
## Year2010         -0.47025      0.22840     -2.06  0.04484 *
## Year2011         -0.49192      0.22828     -2.15  0.03611 *
## Year2012         -0.41768      0.31558     -1.32  0.19180
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.745
## Multiple R-squared:  0.163, Adjusted R-squared:  -0.145
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  8 weights are ~= 1. The remaining 60 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0053 0.9150 0.9610 0.9280 0.9900 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 57.02 1          7.551
## Year              57.02 16          1.135

```

Residuals from first author

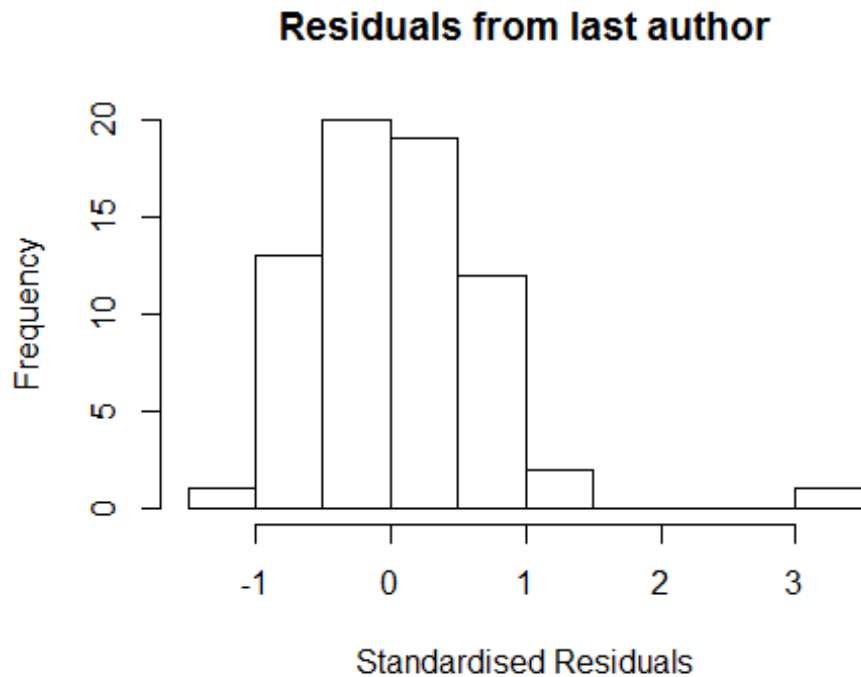


```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 9 0030305457 4.238 1996      1804      3      3.362
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37923 -0.34884  0.00817  0.41414  3.36729
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.87071    0.21814   3.99 0.00021 ***
## FirstAuthorFemale1 0.01978    0.40523   0.05 0.96127
## Year1997      -0.16392    0.34884  -0.47 0.64047
## Year1998      -0.19344    0.35342  -0.55 0.58658
## Year1999       0.06377    0.29822   0.21 0.83155
## Year2000       0.10943    0.48045   0.23 0.82075
## Year2001       0.50852    0.51755   0.98 0.33055
## Year2002       0.03479    0.33002   0.11 0.91646
## Year2003      -0.00353    0.39819  -0.01 0.99296
## Year2004       0.49679    0.24308   2.04 0.04626 *
## Year2005       0.04576    0.47402   0.10 0.92348
## Year2006       0.38076    0.32300   1.18 0.24405
```

```

## Year2007      -0.01971    0.43509   -0.05  0.96405
## Year2008      0.00919    0.37634    0.02  0.98062
## Year2009     -0.14466    0.33473   -0.43  0.66748
## Year2010     -0.46504    0.21826   -2.13  0.03806 *
## Year2011     -0.48671    0.21814   -2.23  0.03018 *
## Year2012     -0.41242    0.30804   -1.34  0.18668
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.729
## Multiple R-squared:  0.154, Adjusted R-squared:  -0.133
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| <= 0.00077 ( < 0.0015);
## 8 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.686  0.921  0.964  0.941  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      1.47e-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 8.245 1          2.871
## Year            8.245 16          1.068

```



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 9 0030305457 4.238 1996      1804      3      3.362
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.37918 -0.36570  0.00817  0.41413  3.36726
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.87074    0.21821   3.99  0.00022 ***
## LastAuthorFemale1 -0.34020    0.21960  -1.55  0.12764
## Year1997      -0.03745    0.24249  -0.15  0.87787
## Year1998      -0.07709    0.34304  -0.22  0.82310
## Year1999       0.06373    0.29838   0.21  0.83174
## Year2000       0.10942    0.48033   0.23  0.82073
## Year2001       0.50844    0.51781   0.98  0.33088
## Year2002       0.03476    0.33007   0.11  0.91655
## Year2003      -0.00357    0.39854  -0.01  0.99289
## Year2004       0.49676    0.24314   2.04  0.04633 *
## Year2005       0.04640    0.48389   0.10  0.92399
## Year2006       0.38489    0.30644   1.26  0.21494
```

```

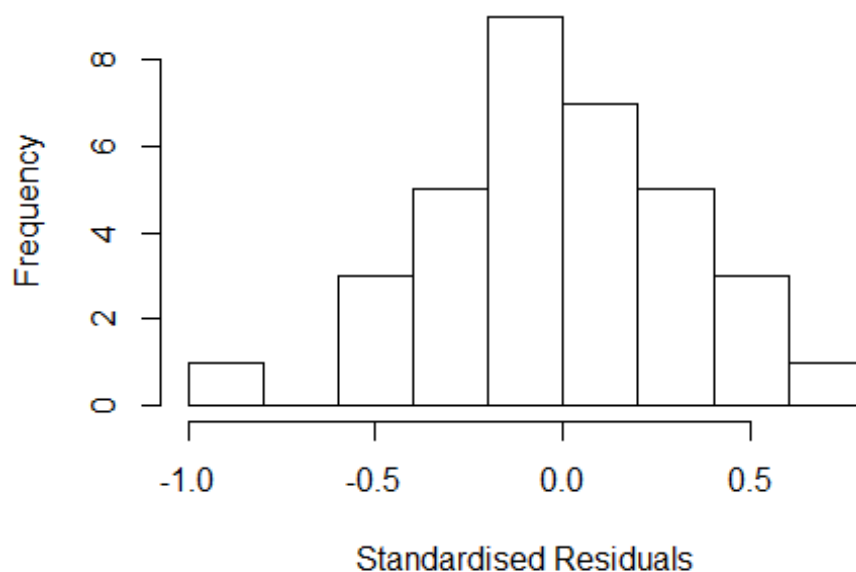
## Year2007          -0.01974      0.43511   -0.05   0.96400
## Year2008           0.00915      0.37656    0.02   0.98071
## Year2009          -0.14468      0.33472   -0.43   0.66742
## Year2010          -0.46507      0.21833   -2.13   0.03810 *
## Year2011          -0.48674      0.21821   -2.23   0.03022 *
## Year2012          -0.41246      0.30811   -1.34   0.18673
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.729
## Multiple R-squared:  0.163, Adjusted R-squared:  -0.122
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## observation 3 is an outlier with |weight| <= 0.0008 ( < 0.0015);
## 7 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.679  0.917  0.965  0.942  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      1.47e-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: 68"
## [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
##    1    4    2    3    3    5    8    2    4    3    4    4    3    2    5
## 2011 2012
##    8    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    2    1    3    1    5    3    0    3    2    2    1    3    1    3

```



```
## 2011 2012
##    5    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    2    1    3    1    5    3    0    3    1    2    1    2    1    3
## 2011 2012
##    3    3
## [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: 1.12246204830937"
## [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 16.53  1          4.066
## Year              16.53 14          1.105
```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -8.43e-01 -2.23e-01 -1.67e-16  1.93e-01  7.68e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.5031    0.1248    4.03  0.00078 ***
## FirstAuthorFemale1 0.2419    0.2385    1.01  0.32400
## Year1998        0.3599    0.1248    2.88  0.00988 **
## Year1999        0.0114    0.2528    0.05  0.96443
## Year2000        0.7539    0.1248    6.04  1e-05 ***
## Year2001        0.8400    0.1947    4.31  0.00042 ***
## Year2002        0.3931    0.1401    2.81  0.01169 *
## Year2004        0.5078    0.1686    3.01  0.00749 **
## Year2005       -0.5031    0.1248   -4.03  0.00078 ***
## Year2006        0.9840    0.3144    3.13  0.00579 **
## Year2007        0.5699    0.1248    4.57  0.00024 ***
## Year2008        0.3819    0.2551    1.50  0.15164
## Year2009        0.2669    0.1248    2.14  0.04642 *
## Year2010        0.3971    0.2550    1.56  0.13677
## Year2011        0.3399    0.5044    0.67  0.50901
## Year2012        0.2166    0.2309    0.94  0.36045
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.54
## Multiple R-squared:  0.467, Adjusted R-squared:  0.0225
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~1. The remaining 22 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.790  0.938  0.961  0.949  0.980  0.996
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.94e-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 12.43  1          3.525
## Year              12.43 14          1.094

## [1] "List of 0 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
## -8.42e-01 -2.23e-01  2.50e-16  1.93e-01  7.69e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.5029     0.1235   4.07 0.00072 ***
## LastAuthorFemale1 0.2421     0.2358   1.03 0.31805
## Year1998          0.3601     0.1235   2.92 0.00924 **
## Year1999          0.0114     0.2516   0.05 0.96429
## Year2000          0.7541     0.1235   6.11 9.1e-06 ***
## Year2001          0.8399     0.1943   4.32 0.00041 ***
## Year2002          0.3932     0.1389   2.83 0.01107 *
## Year2004          0.5079     0.1678   3.03 0.00725 **
## Year2005         -0.5029     0.1235  -4.07 0.00072 ***
## Year2006          0.9840     0.3132   3.14 0.00563 **
## Year2007          0.5701     0.1235   4.62 0.00021 ***
## Year2008          0.3821     0.2540   1.50 0.14989
## Year2009          0.2671     0.1235   2.16 0.04431 *
## Year2010          0.3970     0.2546   1.56 0.13629
## Year2011          0.3392     0.4999   0.68 0.50607
## Year2012          0.2170     0.2289   0.95 0.35562
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.55
## Multiple R-squared: 0.466, Adjusted R-squared: 0.0211
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~ 1. The remaining 22 ones are summarized as
##      Min. 1st Qu. Median   Mean 3rd Qu.    Max.
## 0.798 0.941 0.962 0.951 0.980 0.996
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol

```

```

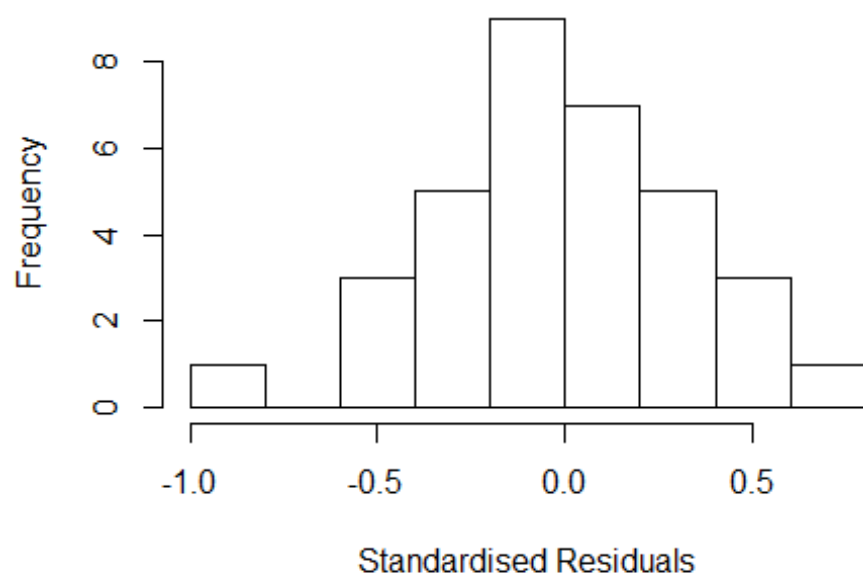
##          1.55e+00          5.00e-01          4.69e+00          1.00e-07
##          rel.tol          solve.tol          eps.outlier          eps.x
##          1.00e-07          1.00e-07          2.94e-03          1.82e-12
## warn.limit.reject warn.limit.meanrw
##          5.00e-01          5.00e-01
## 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: 34"
## [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
##      3      4      3      3      4      7      5      7      8      2      9      3      7      7      9
## 2011 2012
##      7      12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      4      1      2      3      5      3      4      6      1      7      3      7      5      8
## 2011 2012
##      5      11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      1      3      1      1      2      3      3      3      4      1      7      3      7      5      8
## 2011 2012
##      5      11
## [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: 1.25992104989487"
## [1] "Regression 1: First author gender, last author gender, team size,
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

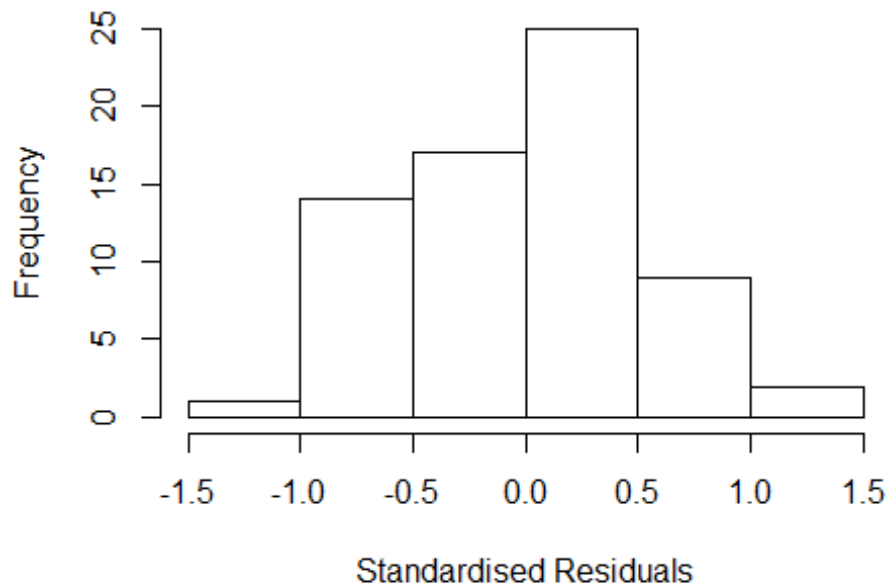
```

Residuals from last author



##		GVIF	Df	$GVIF^{(1/(2*Df))}$
##	FirstAuthorFemale	NaN	1	NaN
##	LastAuthorFemale	NaN	1	NaN
##	UniqueAuthors	NaN	2	NaN
##	Year	NaN	16	NaN

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.0611 -0.4290 0.0463 0.3773 1.2017
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 5.17e-01 2.19e-08 2.36e+07 < 2e-16 ***
## FirstAuthorFemale1 -2.63e-01 3.79e-01 -6.90e-01 0.49077
## LastAuthorFemale1 7.65e-01 3.85e-01 1.99e+00 0.05258 .
## UniqueAuthors2 3.77e-01 2.21e-01 1.70e+00 0.09551 .
## UniqueAuthors3 5.44e-01 1.53e-01 3.56e+00 0.00087 ***
## Year1997 1.84e-01 5.15e-01 3.60e-01 0.72180
## Year1998 3.03e-03 3.82e-01 1.00e-02 0.99370
## Year1999 8.34e-01 8.43e-09 9.89e+07 < 2e-16 ***
## Year2000 3.04e-01 4.42e-01 6.90e-01 0.49509
## Year2001 2.60e-01 1.65e-01 1.57e+00 0.12281
```

```

## Year2002          6.17e-01  1.32e-01  4.68e+00  2.5e-05 ***
## Year2003          9.13e-01  2.51e-01  3.64e+00  0.00068 ***
## Year2004          4.15e-01  5.18e-01  8.00e-01  0.42760
## Year2005          5.52e-01  0.00e+00      Inf  < 2e-16 ***
## Year2006          3.55e-01  1.69e-01  2.10e+00  0.04144 *
## Year2007          1.93e-02  2.43e-01  8.00e-02  0.93711
## Year2008          5.44e-01  2.61e-01  2.08e+00  0.04264 *
## Year2009          3.24e-02  2.45e-01  1.30e-01  0.89566
## Year2010          3.82e-01  2.25e-01  1.69e+00  0.09698 .
## Year2011          4.19e-02  2.31e-01  1.80e-01  0.85671
## Year2012          3.02e-01  2.00e-01  1.52e+00  0.13640
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.64
## Multiple R-squared:  0.217, Adjusted R-squared:  -0.116
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.705  0.928  0.955  0.936  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      1.47e-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 -7.844e+14  1      NaN
## LastAuthorFemale  -7.798e+14  1      NaN
## Year              -5.789e+15 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 = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.1148 -0.4247  0.0214  0.3887  1.1664
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    5.17e-01   4.87e-08  1.06e+07 < 2e-16 ***
## FirstAuthorFemale1 -2.81e-01   2.40e-01 -1.17e+00  0.24758
## LastAuthorFemale1  8.01e-01   2.39e-01  3.35e+00  0.00157 **
## Year1997         1.90e-01   5.04e-01  3.80e-01  0.70770
## Year1998        -1.53e-02   3.55e-01 -4.00e-02  0.96586
## Year1999         8.34e-01   1.06e-08  7.90e+07 < 2e-16 ***
## Year2000         4.93e-01   6.49e-01  7.60e-01  0.45190
## Year2001         3.88e-01   8.90e-02  4.36e+00  6.7e-05 ***
## Year2002         6.17e-01   1.32e-01  4.69e+00  2.2e-05 ***
## Year2003         9.14e-01   2.49e-01  3.66e+00  0.00061 ***
## Year2004         4.82e-01   6.33e-01  7.60e-01  0.44961
## Year2005         5.52e-01   1.93e-08  2.86e+07 < 2e-16 ***
## Year2006         5.25e-01   1.80e-01  2.91e+00  0.00545 **
## Year2007         1.42e-01   3.21e-01  4.40e-01  0.65900
## Year2008         5.98e-01   2.86e-01  2.09e+00  0.04177 *
## Year2009         1.07e-01   2.30e-01  4.70e-01  0.64298
## Year2010         4.95e-01   1.90e-01  2.61e+00  0.01197 *
## Year2011         2.74e-01   2.18e-01  1.26e+00  0.21508
## Year2012         3.38e-01   1.83e-01  1.85e+00  0.07084 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.67
## Multiple R-squared:  0.153, Adjusted R-squared:  -0.158
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 55 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.743  0.901  0.952  0.932  0.984  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      1.47e-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

```

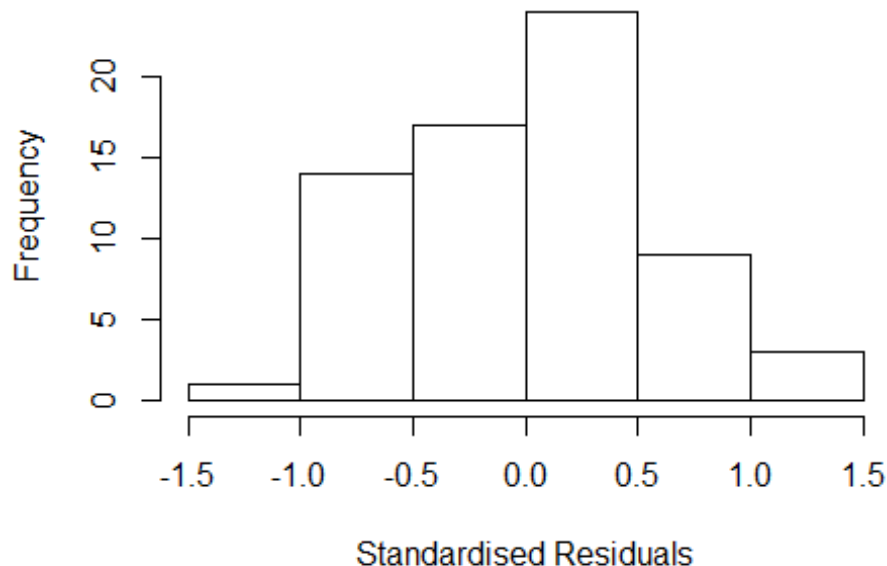


```
##           0           1000           0
##           psi           subsampling           cov
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## 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
```

Residuals from first and last author



```
##           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 = "MM"
```

```

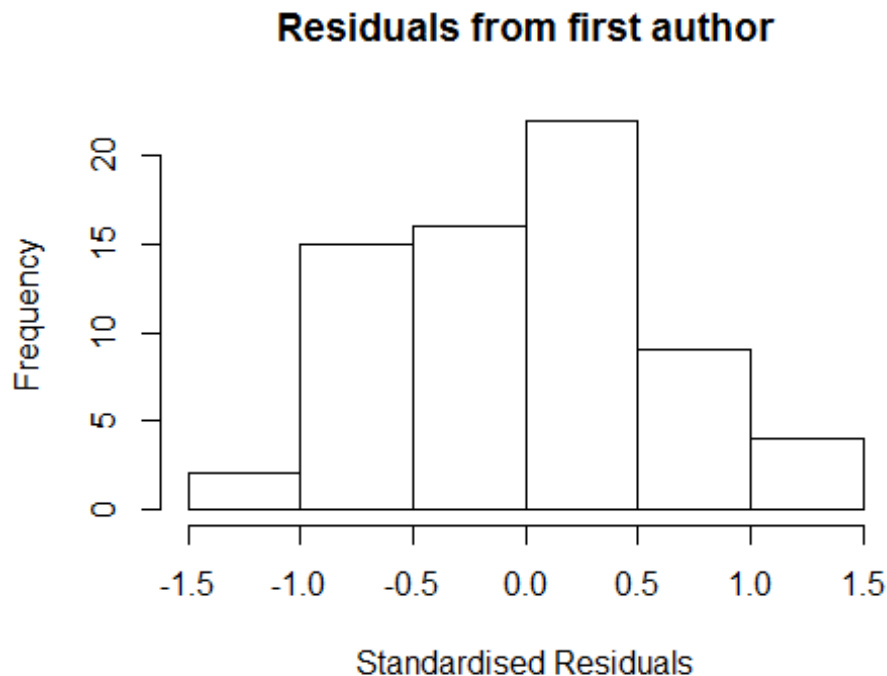
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1306 -0.4691  0.0164  0.3894  1.1473
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    5.17e-01   3.15e-08  1.64e+07 < 2e-16 ***
## FirstAuthorFemale1 3.36e-01   3.65e-01  9.20e-01  0.36083
## Year1997        1.93e-01   4.99e-01  3.90e-01  0.70080
## Year1998        1.69e-01   3.65e-01  4.60e-01  0.64604
## Year1999        8.34e-01   0.00e+00      Inf < 2e-16 ***
## Year2000        4.92e-01   6.41e-01  7.70e-01  0.44566
## Year2001        1.90e-01   2.30e-01  8.30e-01  0.41227
## Year2002        6.17e-01   1.31e-01  4.69e+00  2.1e-05 ***
## Year2003        9.14e-01   2.49e-01  3.67e+00  0.00058 ***
## Year2004        4.84e-01   6.19e-01  7.80e-01  0.43803
## Year2005        5.52e-01   3.49e-08  1.58e+07 < 2e-16 ***
## Year2006        5.54e-01   1.79e-01  3.10e+00  0.00319 **
## Year2007        1.42e-01   3.19e-01  4.50e-01  0.65822
## Year2008        6.14e-01   3.02e-01  2.03e+00  0.04746 *
## Year2009        2.80e-01   2.35e-01  1.19e+00  0.23852
## Year2010        4.94e-01   1.88e-01  2.62e+00  0.01155 *
## Year2011        2.74e-01   2.18e-01  1.26e+00  0.21483
## Year2012        3.57e-01   1.81e-01  1.97e+00  0.05423 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.686
## Multiple R-squared:  0.114, Adjusted R-squared:  -0.187
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
##  10 weights are ~= 1. The remaining 58 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.761  0.908  0.950  0.935  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      1.47e-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

## 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 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 = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1059 -0.4359  0.0229  0.3893  1.1736
##
## Coefficients:
```

```

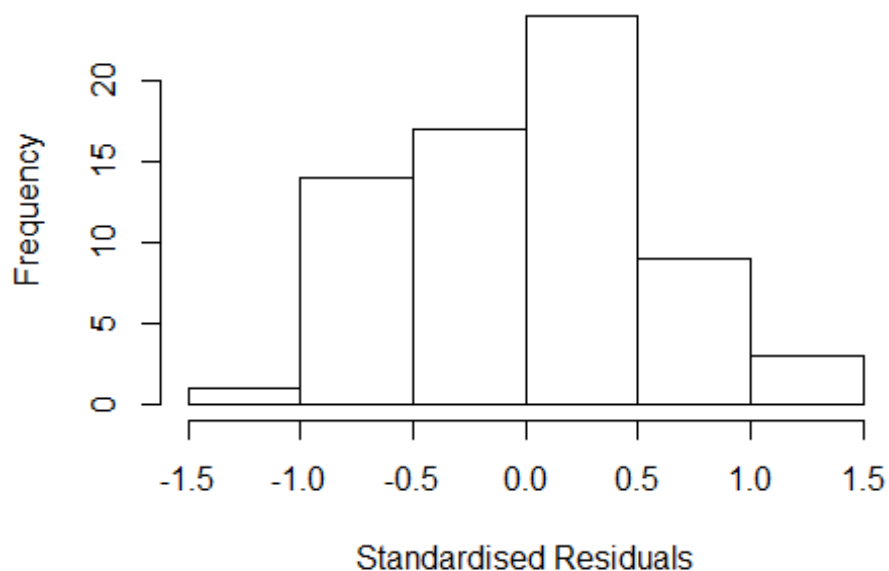
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.5170     0.0000    Inf < 2e-16 ***
## LastAuthorFemale1 0.5972     0.3059     1.95 0.05650 .
## Year1997          0.1923     0.5001     0.38 0.70214
## Year1998         -0.0922     0.3059    -0.30 0.76427
## Year1999          0.8340     0.0000    Inf < 2e-16 ***
## Year2000          0.4925     0.6423     0.77 0.44684
## Year2001          0.2958     0.1296     2.28 0.02680 *
## Year2002          0.6165     0.1315     4.69 2.2e-05 ***
## Year2003          0.9139     0.2489     3.67 0.00059 ***
## Year2004          0.4835     0.6215     0.78 0.44025
## Year2005          0.5520     0.0000    Inf < 2e-16 ***
## Year2006          0.5132     0.1789     2.87 0.00603 **
## Year2007          0.1422     0.3196     0.44 0.65837
## Year2008          0.5889     0.2813     2.09 0.04136 *
## Year2009          0.1526     0.2175     0.70 0.48637
## Year2010          0.4942     0.1887     2.62 0.01163 *
## Year2011          0.2737     0.2178     1.26 0.21486
## Year2012          0.3304     0.1828     1.81 0.07664 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.683
## Multiple R-squared:  0.15, Adjusted R-squared:  -0.14
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 56 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.749  0.912  0.954  0.935  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      1.47e-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: 68"
## [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
##    2    1    2    5    5   10    6    5    8    6    5    7    8    9    6
## 2011 2012
##    6    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    2    0    0    5    2    5    4    4    2    1    4    4    2
## 2011 2012
##    3    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    2    0    0    5    2    3    4    3    2    0    4    3    2
## 2011 2012
##    3    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 1"

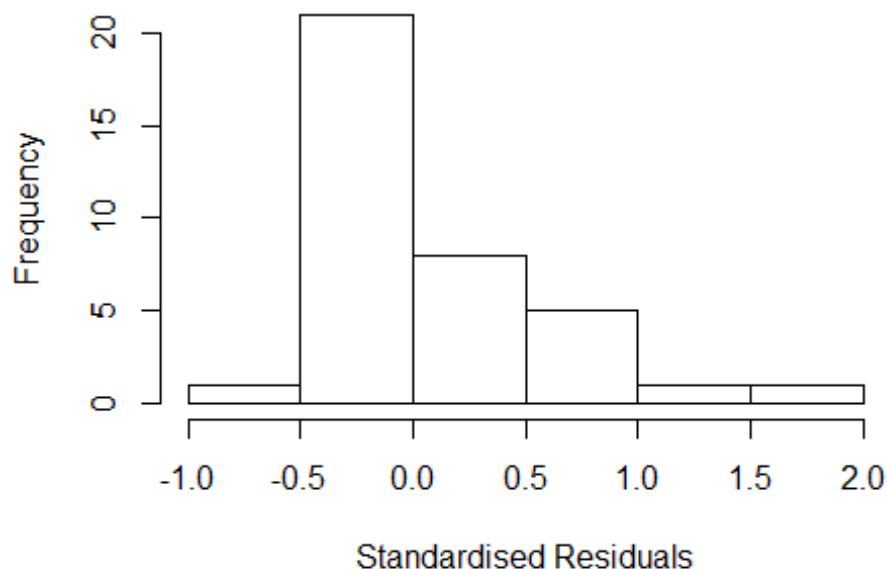
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 1.18920711500272"
## [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"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.519e+14  1      NaN
## LastAuthorFemale  -4.345e+15  1      NaN
## Year              -5.040e+28 12      NaN
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -0.645 -0.319 -0.113  0.259  1.503
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5918     0.2139   7.44 1.9e-07 ***
## FirstAuthorFemale1 -1.2928     0.2139  -6.04 4.4e-06 ***
## LastAuthorFemale1   0.5125     0.0931   5.50 1.6e-05 ***
## Year1998           -0.7808     0.2608  -2.99 0.00668 **
## Year2001           -0.4745     0.0931  -5.09 4.2e-05 ***
## Year2002           -1.1013     0.2563  -4.30 0.00029 ***
## Year2003           -0.5420     0.4268  -1.27 0.21734
## Year2004           -0.8605     0.3676  -2.34 0.02871 *
## Year2005           -0.7623     0.2679  -2.85 0.00942 **
## Year2006           -0.9888     0.3098  -3.19 0.00421 **
## Year2008           -1.1212     0.3614  -3.10 0.00519 **
## Year2009           -0.8325     0.7327  -1.14 0.26813
## Year2010           -0.6123     0.3112  -1.97 0.06189 .
## Year2011           -0.5927     0.3984  -1.49 0.15103
## Year2012           -0.5607     0.4262  -1.32 0.20188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.224, Adjusted R-squared:  -0.27
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 32 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.480  0.933  0.972  0.931  0.988  0.997
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+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-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 8.912 1 2.985
## Year 8.912 12 1.095

## [1] "List of 0 outliers 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
## -0.645 -0.326 -0.113 0.301 1.496
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08e+00 1.92e-01 5.61 1e-05 ***
## FirstAuthorFemale1 -7.80e-01 1.92e-01 -4.06 0.00049 ***
## Year1998 -2.68e-01 2.43e-01 -1.10 0.28137
## Year2001 3.80e-02 5.63e-08 674773.50 < 2e-16 ***
## Year2002 -5.89e-01 2.39e-01 -2.47 0.02142 *
## Year2003 -2.89e-02 4.15e-01 -0.07 0.94510
## Year2004 -3.47e-01 3.55e-01 -0.98 0.33835
## Year2005 -8.03e-02 2.48e-01 -0.32 0.74910
## Year2006 -4.76e-01 2.95e-01 -1.61 0.12008
## Year2008 -6.09e-01 3.48e-01 -1.75 0.09405 .
## Year2009 -3.14e-01 7.23e-01 -0.43 0.66842
## Year2010 -9.99e-02 2.97e-01 -0.34 0.73931
## Year2011 -8.04e-02 3.86e-01 -0.21 0.83700
## Year2012 -4.76e-02 4.15e-01 -0.11 0.90961
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.583
## Multiple R-squared: 0.205, Adjusted R-squared: -0.244
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 32 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.490 0.934 0.970 0.931 0.982 0.997
## Algorithmic parameters:
## tuning.chi bb tuning.psi refine.tol
## 1.55e+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-03 1.82e-12
## warn.limit.reject warn.limit.meanrw
## 5.00e-01 5.00e-01

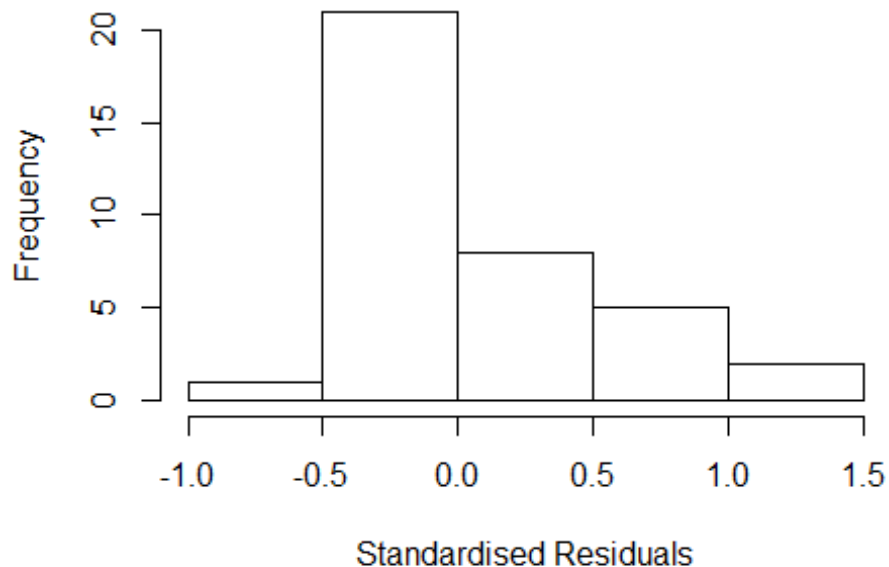
```



```
##      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"

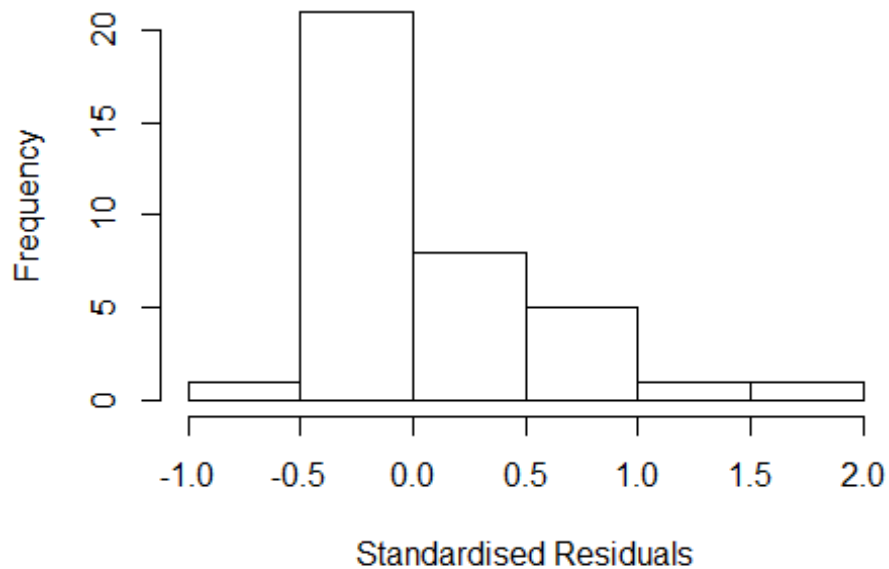
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```

Residuals from first author



```
##              GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.146 1      1.071
## Year             1.146 12     1.006
```

Residuals from last author



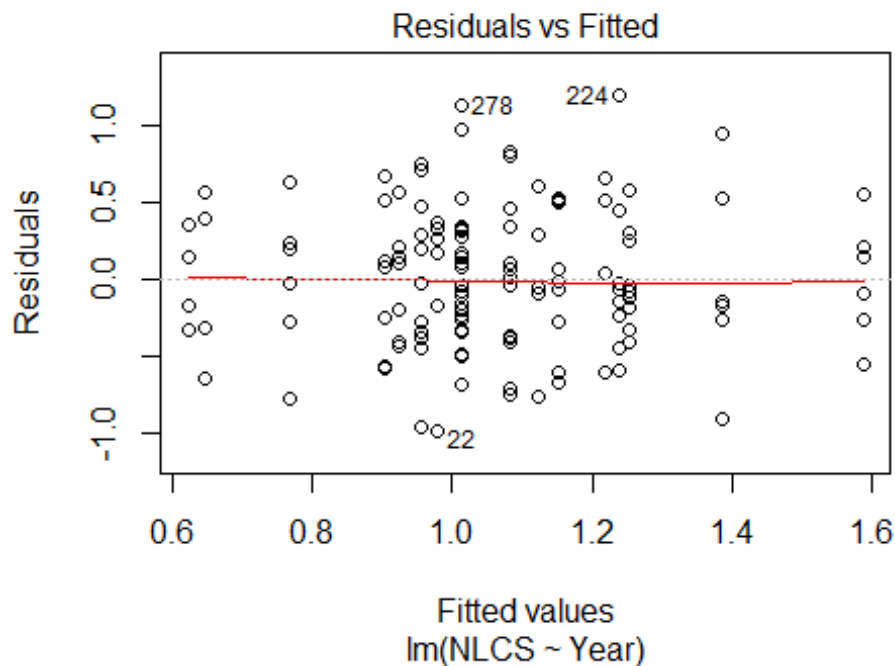
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.645 -0.359 -0.166 0.301 1.448
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.299 0.000 Inf <2e-16 ***
## LastAuthorFemale1 -0.187 0.397 -0.47 0.6422
## Year1998 0.512 0.149 3.44 0.0022 **
## Year2001 0.696 0.194 3.59 0.0015 **
## Year2002 0.191 0.141 1.36 0.1869
## Year2003 0.757 0.367 2.06 0.0505 .
## Year2004 0.442 0.306 1.44 0.1622
## Year2005 0.759 0.205 3.71 0.0012 **
## Year2006 0.304 0.222 1.37 0.1849
## Year2008 0.174 0.289 0.60 0.5532
## Year2009 0.515 0.716 0.72 0.4792
## Year2010 0.680 0.224 3.03 0.0059 **
```

```

## Year2011          0.700      0.329      2.13      0.0443 *
## Year2012          0.738      0.366      2.02      0.0555 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.622
## Multiple R-squared:  0.159, Adjusted R-squared:  -0.317
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.567  0.946  0.969  0.940  0.984  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      2.70e-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: 37"
## [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
##   10   14   16   11   10   11   10   10   14   10   16   16   13   24   25
## 2011 2012
##   25   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6    6    4    6    7    4    6    9    5    9    9    7   12   13
## 2011 2012
##   13   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    5    5    3    5    7    4    6    7    4    9    5    7   11   12

```

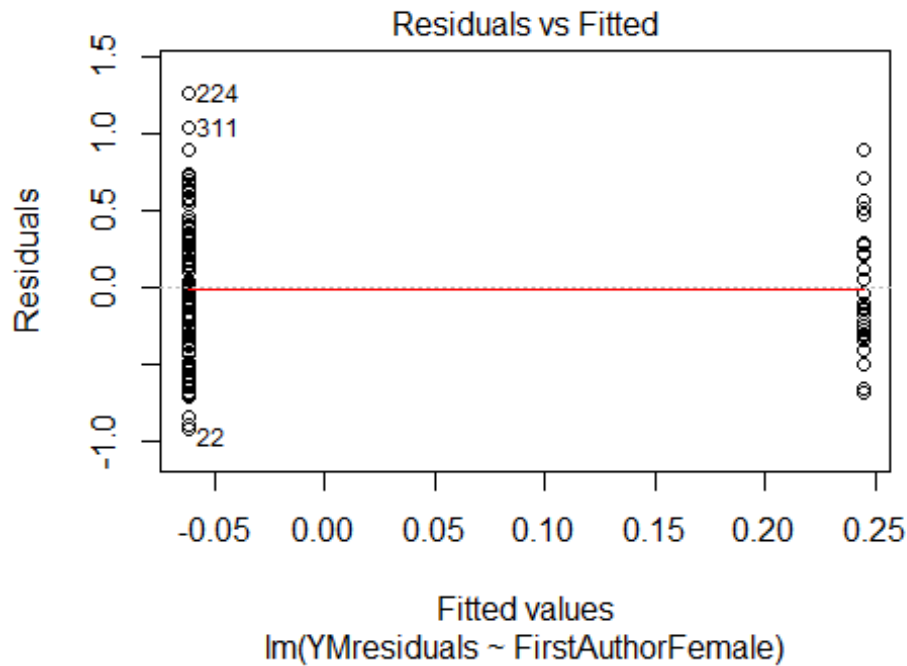
```
## 2011 2012
## 11 9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 6.5, df = 16, p-value = 1
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.29, df = 1, p-value = 0.6
## [1] "Female first author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male first author team size 2018 geometric mean: 2.70192007704123"
## 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: 2.28942848510666"
## [1] "Male last author team size 2018 geometric mean: 2.91295063024394"

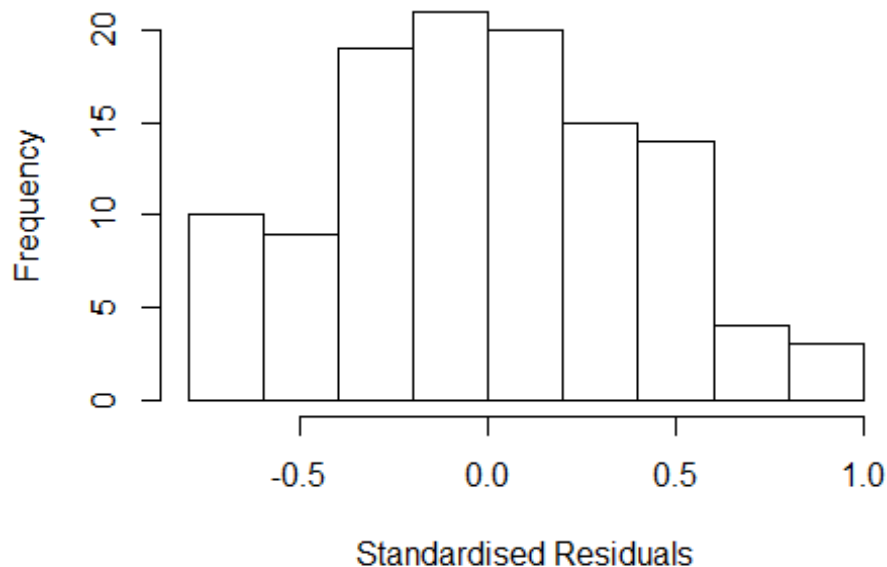
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, 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	4.072	1	2.018
LastAuthorFemale	1.951	1	1.397
UniqueAuthors	9.374	4	1.323
Year	19.004	16	1.096

Residuals from first and last author and team size



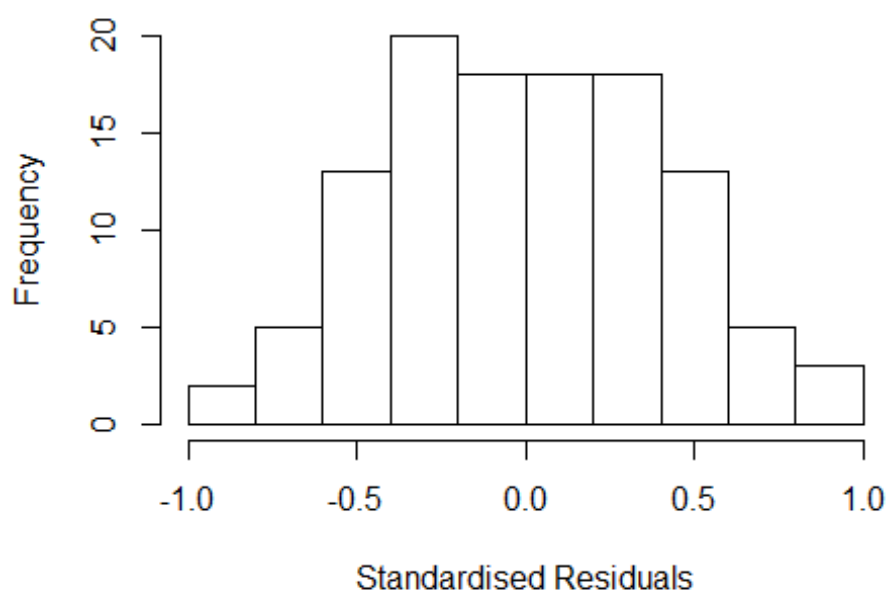
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7915 -0.2671 -0.0217 0.3186 0.9413
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1431 0.2199 5.20 1.2e-06 ***
## FirstAuthorFemale1 0.3562 0.1589 2.24 0.027 *
## LastAuthorFemale1 -0.2146 0.1509 -1.42 0.159
## UniqueAuthors2 0.1243 0.1137 1.09 0.277
## UniqueAuthors3 0.1512 0.1505 1.00 0.318
## UniqueAuthors4 -0.0274 0.2244 -0.12 0.903
## UniqueAuthors5 0.3121 0.2469 1.26 0.209
## Year1997 -0.3515 0.3833 -0.92 0.361
## Year1998 0.5352 0.2462 2.17 0.032 *
## Year1999 -0.7193 0.2862 -2.51 0.014 *
```

```

## Year2000          -0.2934      0.2795    -1.05      0.297
## Year2001          -0.2693      0.3121    -0.86      0.390
## Year2002          -0.5708      0.3629    -1.57      0.119
## Year2003          -0.0056      0.3110    -0.02      0.986
## Year2004          -0.2989      0.2791    -1.07      0.287
## Year2005           0.0906      0.4613      0.20      0.845
## Year2006          -0.0893      0.2601    -0.34      0.732
## Year2007          -0.0338      0.2773    -0.12      0.903
## Year2008          -0.3625      0.2674    -1.36      0.178
## Year2009          -0.3100      0.2663    -1.16      0.247
## Year2010          -0.2506      0.2645    -0.95      0.346
## Year2011          -0.2975      0.3267    -0.91      0.365
## Year2012          -0.4043      0.2837    -1.42      0.158
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.452
## Multiple R-squared:  0.302, Adjusted R-squared:  0.135
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 110 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.644  0.903  0.958  0.928  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      8.70e-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 3.017 1      1.737
## LastAuthorFemale  1.622 1      1.274
## Year              3.264 16      1.038

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -0.8952 -0.2666 -0.0503 0.3075 0.9992
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1411 0.2185 5.22 1e-06 ***
## FirstAuthorFemale1 0.3954 0.1414 2.80 0.0062 **
## LastAuthorFemale1 -0.2814 0.1335 -2.11 0.0376 *
## Year1997 -0.2458 0.3695 -0.67 0.5075
## Year1998 0.5540 0.2542 2.18 0.0317 *
## Year1999 -0.6296 0.2547 -2.47 0.0152 *
## Year2000 -0.2224 0.2596 -0.86 0.3937
## Year2001 -0.2407 0.2930 -0.82 0.4134
## Year2002 -0.4853 0.3765 -1.29 0.2006
## Year2003 0.1825 0.2913 0.63 0.5326
## Year2004 -0.1869 0.2815 -0.66 0.5085
## Year2005 0.2621 0.3587 0.73 0.4667
```

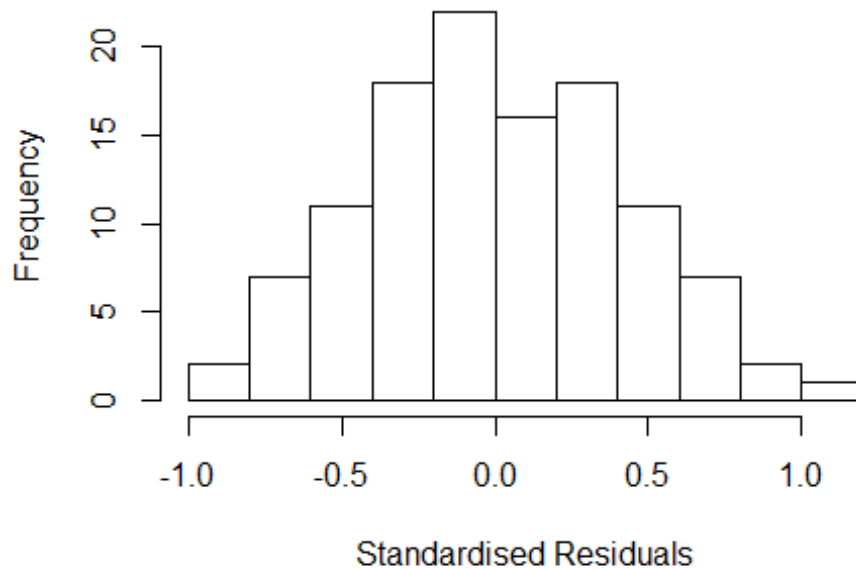


```

## Year2006          0.0473      0.2477      0.19      0.8491
## Year2007          0.0671      0.2598      0.26      0.7968
## Year2008         -0.2398      0.2495     -0.96      0.3389
## Year2009         -0.2066      0.2486     -0.83      0.4080
## Year2010         -0.1553      0.2519     -0.62      0.5391
## Year2011         -0.2206      0.2984     -0.74      0.4615
## Year2012         -0.2989      0.2744     -1.09      0.2788
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.472
## Multiple R-squared:  0.284, Adjusted R-squared:  0.149
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 111 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.633  0.909  0.960  0.933  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      8.70e-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 2.391 1      1.546
## Year              2.391 16      1.028

```

Residuals from first author



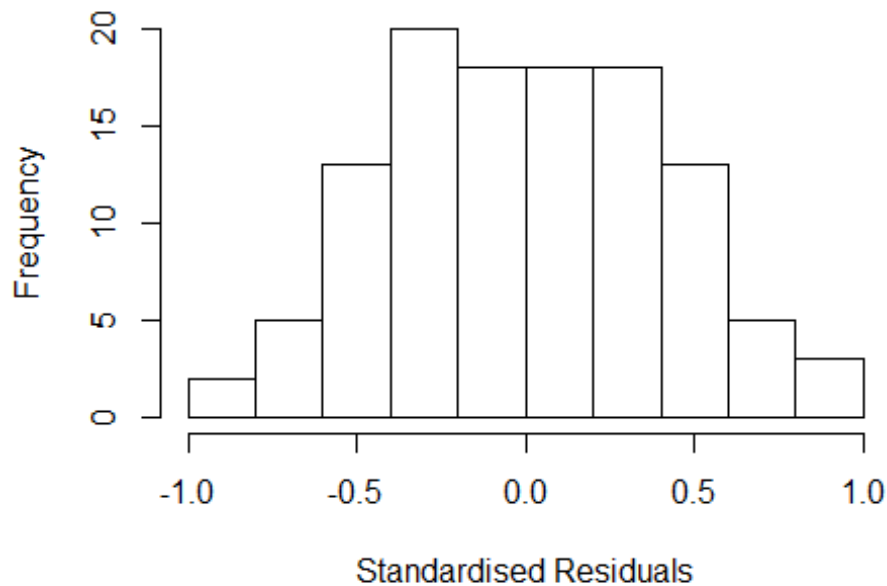
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9093 -0.2948 -0.0455 0.3125 1.0162
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1408 0.2183 5.23 1e-06 ***
## FirstAuthorFemale1 0.3295 0.1345 2.45 0.0161 *
## Year1997 -0.2315 0.3584 -0.65 0.5198
## Year1998 0.5543 0.2540 2.18 0.0315 *
## Year1999 -0.6805 0.2525 -2.69 0.0083 **
## Year2000 -0.2221 0.2594 -0.86 0.3940
## Year2001 -0.2404 0.2936 -0.82 0.4150
## Year2002 -0.4851 0.3765 -1.29 0.2007
## Year2003 0.1427 0.2926 0.49 0.6269
## Year2004 -0.1658 0.2816 -0.59 0.5575
## Year2005 0.2619 0.3581 0.73 0.4664
## Year2006 0.0284 0.2446 0.12 0.9077
```

```

## Year2007          -0.0321      0.2367   -0.14   0.8926
## Year2008          -0.2712      0.2550   -1.06   0.2902
## Year2009          -0.2442      0.2560   -0.95   0.3425
## Year2010          -0.1720      0.2548   -0.68   0.5013
## Year2011          -0.1962      0.3007   -0.65   0.5157
## Year2012          -0.3392      0.2740   -1.24   0.2186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.475
## Multiple R-squared:  0.255, Adjusted R-squared:  0.125
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 111 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.626  0.906   0.960   0.931   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      8.70e-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 1.43 1          1.196
## Year              1.43 16          1.011

```

Residuals from last author



```
## [1] "List of 0 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
## -0.9812 -0.2912 -0.0142 0.3199 1.1484
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1380 0.2163 5.26 8.6e-07 ***
## LastAuthorFemale1 -0.1566 0.1394 -1.12 0.264
## Year1997 -0.1567 0.3137 -0.50 0.618
## Year1998 0.5576 0.2519 2.21 0.029 *
## Year1999 -0.3975 0.2953 -1.35 0.181
## Year2000 -0.2190 0.2573 -0.85 0.397
## Year2001 -0.2369 0.2895 -0.82 0.415
## Year2002 -0.4835 0.3668 -1.32 0.190
## Year2003 0.2679 0.3555 0.75 0.453
## Year2004 -0.0619 0.2868 -0.22 0.830
## Year2005 0.2593 0.3525 0.74 0.464
## Year2006 0.1265 0.2439 0.52 0.605
```

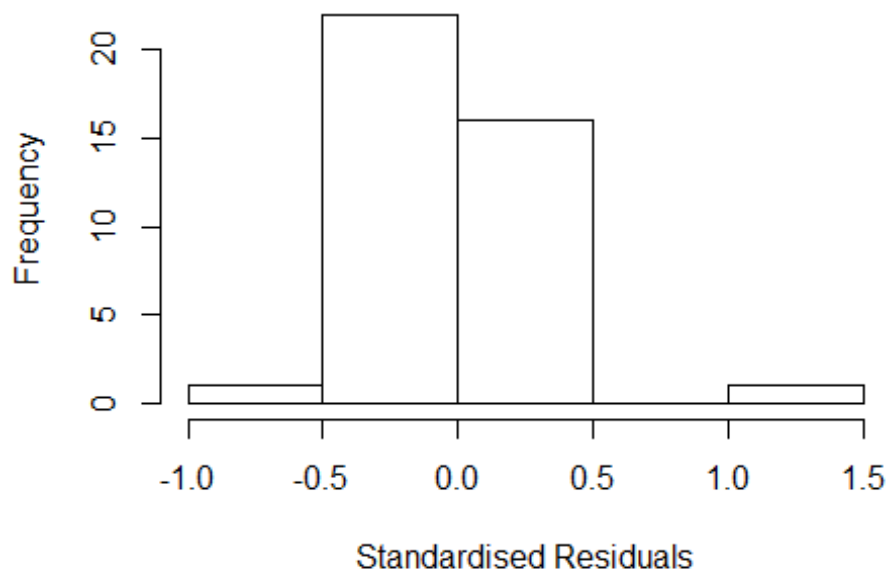
```

## Year2007          0.0882      0.2734      0.32      0.748
## Year2008          -0.1930      0.2493     -0.77      0.441
## Year2009          -0.1404      0.2647     -0.53      0.597
## Year2010          -0.1229      0.2482     -0.50      0.622
## Year2011           0.0583      0.2619      0.22      0.824
## Year2012          -0.2789      0.2567     -1.09      0.280
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.51
## Multiple R-squared:  0.21,   Adjusted R-squared:  0.0715
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 106 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.591  0.902  0.961  0.933  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      8.70e-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: 115"
## [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
##    4    2    6    3   10    3    4    6    2    4    8    2    4    8    3
## 2011 2012
##    5    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    0    3    2    1    3    2    2    4    4    1    3    4    2
## 2011 2012

```

```
##      3      6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      0      3      1      1      2      2      2      4      4      0      3      4      2
## 2011 2012
##      3      6
## [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: 1.31607401295249"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.311e+15 1      3.621e+07
## LastAuthorFemale -1.480e+15 1           NaN
## UniqueAuthors    1.448e+17 2      1.951e+04
## Year              3.413e+32 14      1.452e+01
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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,
```

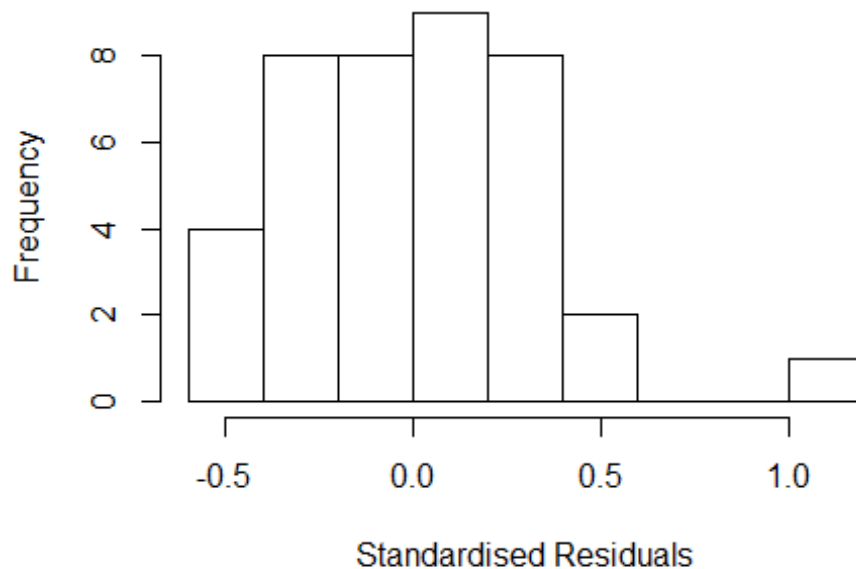
```

##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -6.67e-01 -2.11e-01  5.55e-16  1.80e-01  1.33e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.18e+00   2.10e-08  5.63e+07 < 2e-16 ***
## FirstAuthorFemale1  1.03e+00   3.05e-01   3.38e+00  0.0028 **
## LastAuthorFemale1 -3.91e-01   3.93e-01  -1.00e+00  0.3302
## UniqueAuthors2    1.52e-01   2.54e-01   6.00e-01  0.5553
## UniqueAuthors3    4.96e-01   2.34e-08  2.12e+07 < 2e-16 ***
## Year1997         -1.13e+00   5.34e-01  -2.11e+00  0.0470 *
## Year1999         -7.13e-01   1.22e-01  -5.84e+00  8.4e-06 ***
## Year2000         -2.68e-01   1.59e-08  -1.69e+07 < 2e-16 ***
## Year2001         -5.16e-01   2.26e-08  -2.28e+07 < 2e-16 ***
## Year2002         -1.12e+00   3.44e-01  -3.27e+00  0.0037 **
## Year2003         -3.69e-01   3.42e-01  -1.08e+00  0.2923
## Year2004         -5.26e-01   2.99e-01  -1.76e+00  0.0926 .
## Year2005         -6.66e-01   1.81e-01  -3.68e+00  0.0014 **
## Year2006         -6.24e-01   4.82e-01  -1.29e+00  0.2098
## Year2008         -3.89e-01   2.64e-01  -1.47e+00  0.1557
## Year2009         -3.01e-01   1.51e-01  -2.00e+00  0.0587 .
## Year2010         -3.29e-01   4.91e-01  -6.70e-01  0.5098
## Year2011         -5.67e-01   1.17e-01  -4.84e+00  8.8e-05 ***
## Year2012         -5.51e-01   7.95e-02  -6.93e+00  7.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.346
## Multiple R-squared:  0.497, Adjusted R-squared:  0.0654
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 32 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.109  0.905  0.949   0.913  0.985   0.996
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.50e-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 -5.148e+14  1      NaN
## LastAuthorFemale -3.934e+14  1      NaN
## Year              -8.028e+15 14      NaN
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -0.5287 -0.2540  0.0119  0.2440  1.0323
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.429      0.186   7.68 8.5e-08 ***
```

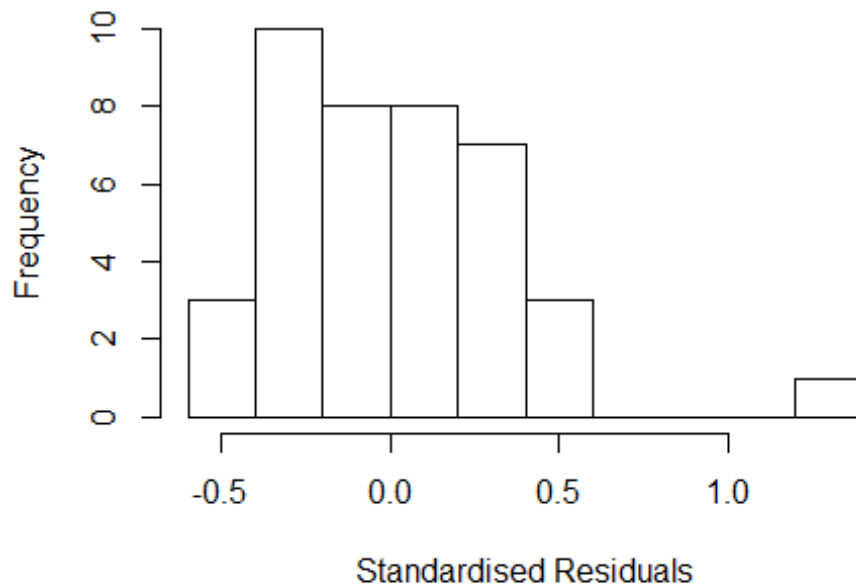


```

## FirstAuthorFemale1    0.840    0.386    2.18  0.04015 *
## LastAuthorFemale1    -0.399    0.337   -1.18  0.24879
## Year1997              -1.176    0.668   -1.76  0.09132 .
## Year1999              -0.960    0.221   -4.34  0.00024 ***
## Year2000              -0.516    0.186   -2.77  0.01080 *
## Year2001              -0.764    0.186   -4.11  0.00043 ***
## Year2002              -1.123    0.358   -3.14  0.00463 **
## Year2003              -0.541    0.284   -1.91  0.06934 .
## Year2004              -0.698    0.261   -2.68  0.01346 *
## Year2005              -0.910    0.243   -3.74  0.00107 **
## Year2006              -0.577    0.733   -0.79  0.43887
## Year2008              -0.573    0.300   -1.91  0.06853 .
## Year2009              -0.517    0.259   -1.99  0.05818 .
## Year2010              -0.501    0.540   -0.93  0.36327
## Year2011              -0.816    0.219   -3.74  0.00108 **
## Year2012              -0.747    0.208   -3.59  0.00155 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.397, Adjusted R-squared:  -0.0217
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 36 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.570  0.956  0.969  0.951  0.992  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      2.50e-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 -2.082e+16  1      NaN
## Year              -2.082e+16 14      NaN

```

Residuals from first author



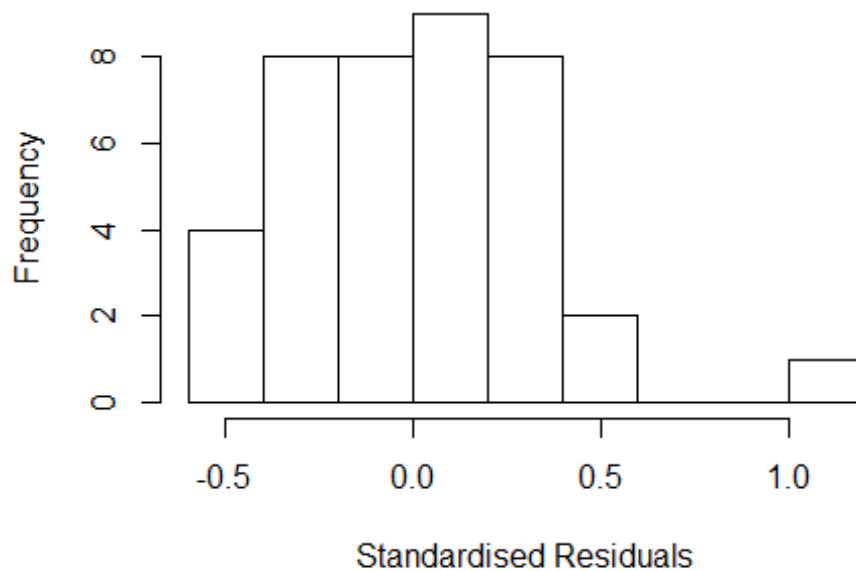
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -5.24e-01 -2.29e-01 6.11e-16 2.45e-01 1.36e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.429 0.187 7.65 6.9e-08 ***
## FirstAuthorFemale1 0.794 0.364 2.18 0.03911 *
## Year1997 -1.530 0.409 -3.74 0.00101 **
## Year1999 -0.960 0.222 -4.33 0.00023 ***
## Year2000 -0.516 0.187 -2.76 0.01083 *
## Year2001 -0.764 0.187 -4.09 0.00042 ***
## Year2002 -1.100 0.383 -2.87 0.00837 **
## Year2003 -0.541 0.285 -1.90 0.07008 .
## Year2004 -0.698 0.262 -2.67 0.01351 *
## Year2005 -0.909 0.244 -3.72 0.00106 **
## Year2006 -0.909 0.401 -2.27 0.03261 *
## Year2008 -0.722 0.278 -2.59 0.01595 *
```

```

## Year2009          -0.517      0.260    -1.99  0.05845 .
## Year2010          -0.501      0.553    -0.91  0.37421
## Year2011          -0.816      0.219    -3.72  0.00106 **
## Year2012          -0.815      0.203    -4.01  0.00052 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.431
## Multiple R-squared:  0.384, Adjusted R-squared:  -0.000802
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 35 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.296  0.948  0.970  0.940  0.987  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      2.50e-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 2.924e+16 1      1.710e+08
## Year              2.924e+16 14      3.873e+00

```

Residuals from last author



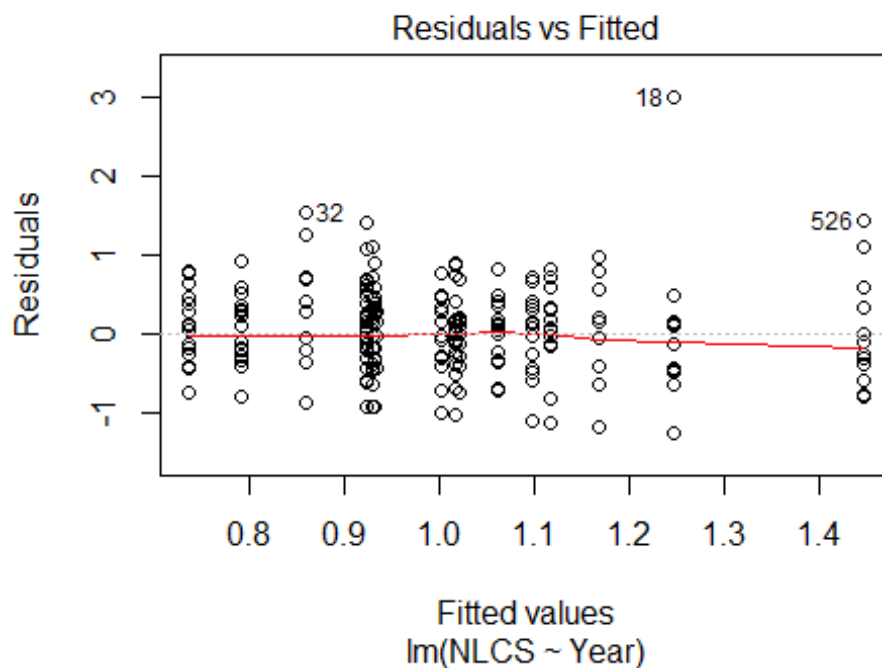
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -7.88e-01 -2.45e-01 5.55e-16 2.45e-01 7.73e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.429 0.185 7.72 5.8e-08 ***
## LastAuthorFemale1 -0.309 0.342 -0.91 0.37383
## Year1997 -0.427 0.388 -1.10 0.28303
## Year1999 -0.959 0.220 -4.36 0.00021 ***
## Year2000 -0.516 0.185 -2.79 0.01018 *
## Year2001 -0.764 0.185 -4.13 0.00038 ***
## Year2002 -0.703 1.034 -0.68 0.50318
## Year2003 -0.541 0.282 -1.92 0.06714 .
## Year2004 -0.698 0.259 -2.69 0.01274 *
## Year2005 -0.912 0.242 -3.76 0.00096 ***
## Year2006 -0.318 0.612 -0.52 0.60825
## Year2008 -0.607 0.299 -2.03 0.05403 .
```

```

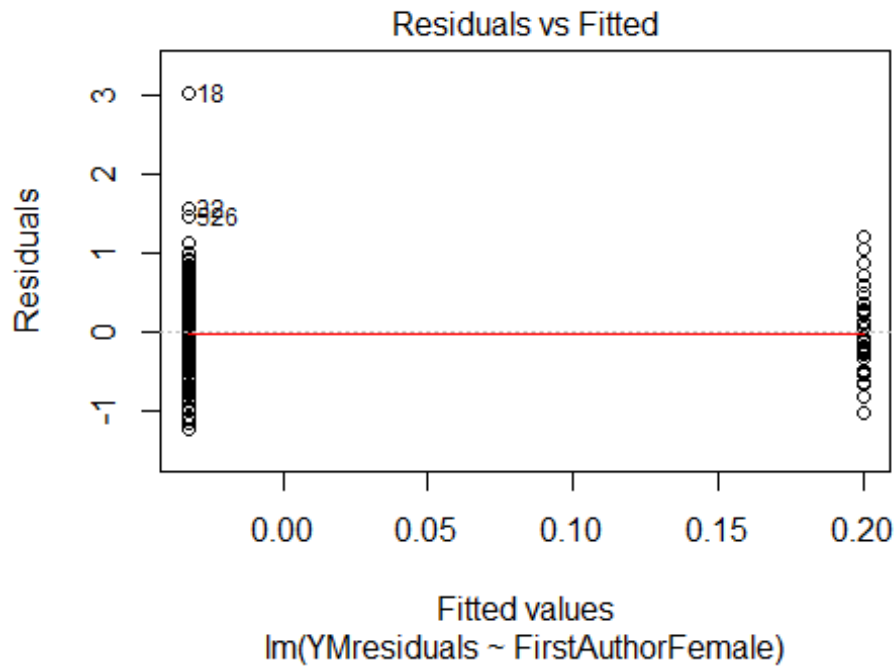
## Year2009          -0.516      0.258   -2.00  0.05669 .
## Year2010          -0.501      0.524   -0.96  0.34824
## Year2011          -0.816      0.217   -3.75  0.00098 ***
## Year2012          -0.763      0.207   -3.68  0.00118 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.467
## Multiple R-squared:  0.299, Adjusted R-squared:  -0.14
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.757  0.936  0.971  0.944  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      2.50e-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: 40"
## [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
##   15   24   20   20   25   39   20   21   25   22    9   31   26   32   31
## 2011 2012
##   31   29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12   14   10   15   15   20   15   11   17   12    4   15   15   19   14
## 2011 2012
##   16   18
##

```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    12    14     8    11    14    18    15    11    14    12     4    10    15    14    14
## 2011 2012
##    15    14
## [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
```

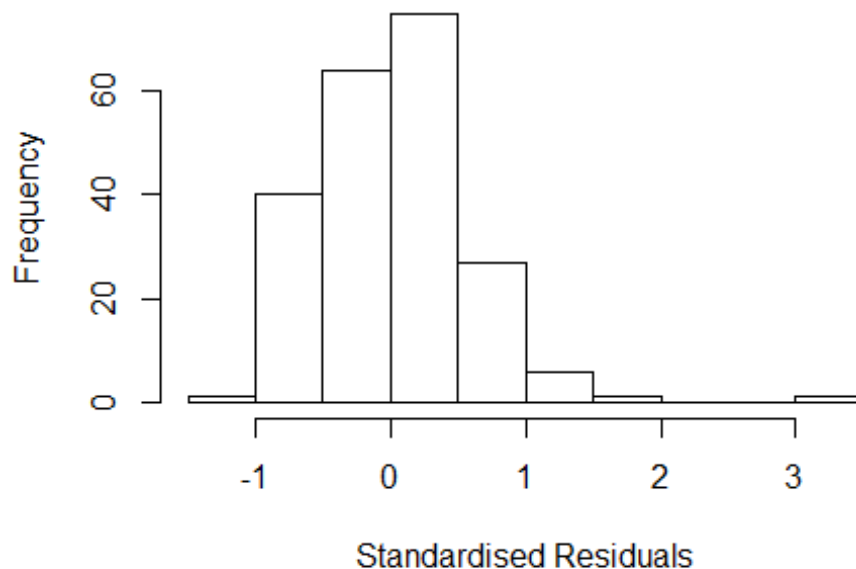


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.21, df = 1, p-value = 0.6
```



```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 1.8612097182042"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.498 1          1.224
## LastAuthorFemale  1.354 1          1.164
## UniqueAuthors    4.706 4          1.214
## Year              6.186 16         1.059
```

Residuals from first and last author and team size



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 18 0030305457 4.238 1996      1804      3      3.212
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1122 -0.3711  0.0437  0.3982  3.2120
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.887695   0.169541   5.24 4.3e-07 ***
## FirstAuthorFemale1 0.205842   0.127745   1.61 0.10875
## LastAuthorFemale1 0.123670   0.121027   1.02 0.30815
## UniqueAuthors2    0.138348   0.087520   1.58 0.11558
## UniqueAuthors3    0.388675   0.113072   3.44 0.00072 ***
## UniqueAuthors4    0.565921   0.338700   1.67 0.09638 .
## UniqueAuthors5    0.312495   0.265879   1.18 0.24132
## Year1997         -0.181433   0.289513  -0.63 0.53161
## Year1998          0.224543   0.333625   0.67 0.50173
## Year1999          0.091363   0.234663   0.39 0.69746
```

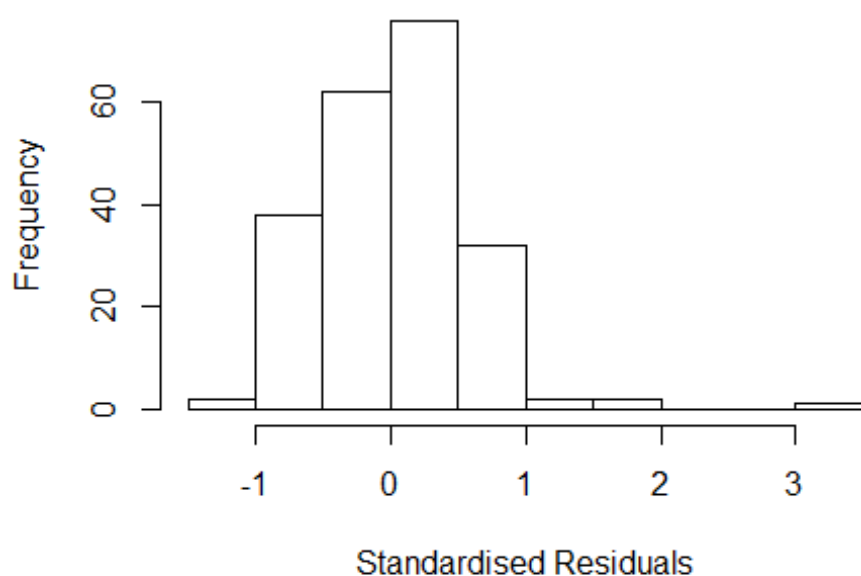


```

## Year2000      -0.167990    0.224768    -0.75    0.45574
## Year2001      -0.210604    0.203436    -1.04    0.30186
## Year2002      -0.094891    0.222199    -0.43    0.66982
## Year2003      -0.045644    0.191994    -0.24    0.81234
## Year2004      -0.177896    0.221240    -0.80    0.42234
## Year2005      -0.146078    0.286752    -0.51    0.61104
## Year2006      -0.150709    0.181808    -0.83    0.40817
## Year2007      -0.000847    0.260513    0.00    0.99741
## Year2008      -0.073245    0.218440    -0.34    0.73776
## Year2009      -0.089650    0.221045    -0.41    0.68551
## Year2010       0.371264    0.238192    1.56    0.12072
## Year2011      -0.048311    0.187442    -0.26    0.79688
## Year2012      -0.186608    0.196725    -0.95    0.34403
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.558
## Multiple R-squared:  0.166, Adjusted R-squared:  0.0705
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.00047);
## 10 weights are ~= 1. The remaining 204 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.336  0.903   0.954   0.925   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           4.65e-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 1.479 1 1.216
## LastAuthorFemale 1.218 1 1.104
## Year 1.794 16 1.018

```

Residuals from first and last author

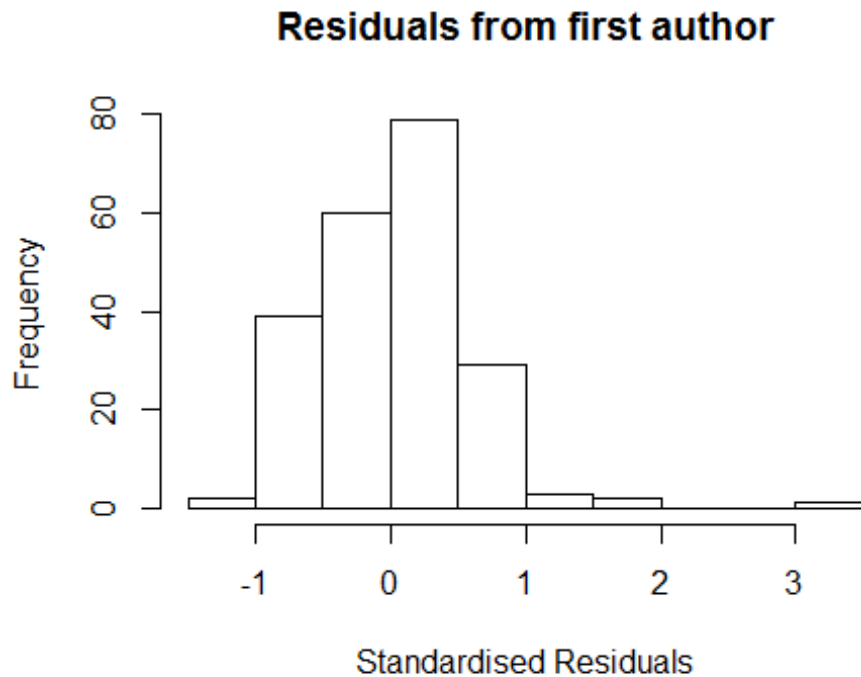


```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 18 0030305457 4.238 1996      1804      3      3.337
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2357 -0.3895  0.0525  0.3798  3.3371
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9009    0.1719   5.24 4.1e-07 ***
## FirstAuthorFemale1 0.2451    0.1277   1.92  0.056 .
## LastAuthorFemale1 0.1778    0.1275   1.40  0.165
## Year1997       -0.1498    0.3015  -0.50  0.620
## Year1998        0.3348    0.3227   1.04  0.301
## Year1999        0.1266    0.2399   0.53  0.598
## Year2000       -0.0754    0.2348  -0.32  0.748
## Year2001       -0.1550    0.2047  -0.76  0.450
## Year2002        0.0336    0.2129   0.16  0.875
## Year2003        0.0531    0.2044   0.26  0.795
## Year2004       -0.0600    0.2265  -0.26  0.791
## Year2005       -0.0132    0.2411  -0.05  0.956
```

```

## Year2006          -0.0106      0.2151   -0.05    0.961
## Year2007           0.0325      0.2632    0.12    0.902
## Year2008          -0.0162      0.2342   -0.07    0.945
## Year2009           0.0592      0.2198    0.27    0.788
## Year2010           0.4696      0.2408    1.95    0.053 .
## Year2011           0.0759      0.2025    0.37    0.708
## Year2012          -0.1324      0.2003   -0.66    0.509
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.106, Adjusted R-squared:  0.0238
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.00047);
## 15 weights are ~= 1. The remaining 199 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.397  0.893   0.952   0.924   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      4.65e-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 1.449 1          1.204
## Year              1.449 16          1.012

```



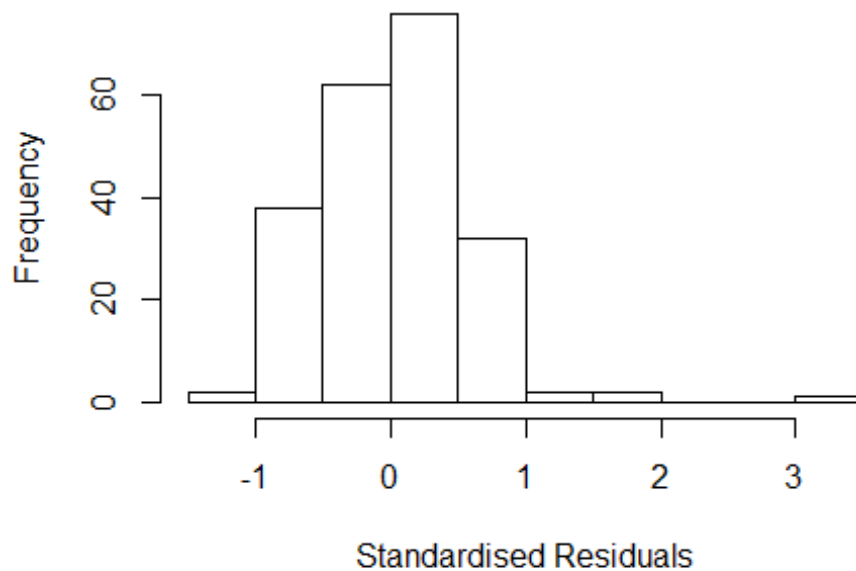
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 18 0030305457 4.238 1996      1804      3      3.337
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2302 -0.3843  0.0418  0.3854  3.3112
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.92678    0.16093   5.76 3.2e-08 ***
## FirstAuthorFemale1 0.28428    0.13122   2.17  0.031 *
## Year1997      -0.15913    0.31242  -0.51  0.611
## Year1998       0.30346    0.31416   0.97  0.335
## Year1999       0.09742    0.23061   0.42  0.673
## Year2000      -0.08497    0.22423  -0.38  0.705
## Year2001      -0.18091    0.19569  -0.92  0.356
## Year2002       0.02628    0.20100   0.13  0.896
## Year2003       0.05195    0.19192   0.27  0.787
## Year2004      -0.07215    0.22250  -0.32  0.746
## Year2005      -0.00757    0.22993  -0.03  0.974
## Year2006       0.01023    0.21105   0.05  0.961
```

```

## Year2007          0.02123      0.26370      0.08      0.936
## Year2008          -0.03110      0.22903     -0.14      0.892
## Year2009          0.03014      0.21126      0.14      0.887
## Year2010          0.44287      0.23293      1.90      0.059 .
## Year2011          0.05535      0.19552      0.28      0.777
## Year2012          -0.13521      0.19087     -0.71      0.480
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.577
## Multiple R-squared:  0.0981, Adjusted R-squared:  0.0202
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.00047);
## 17 weights are ~= 1. The remaining 197 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.402  0.889   0.953   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      4.65e-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

```

Residuals from last author



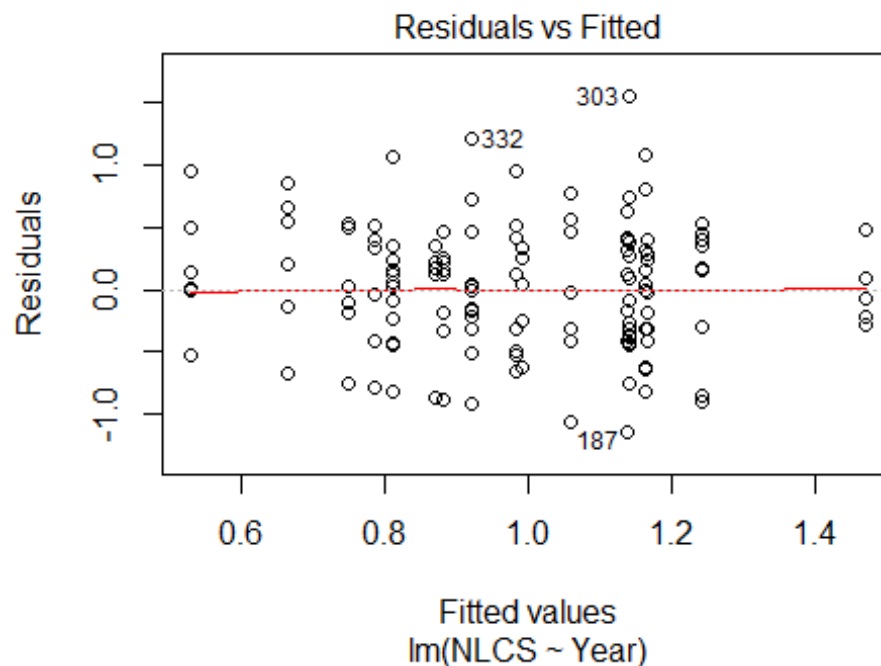
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 18 0030305457 4.238 1996      1804      3      3.337
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2693 -0.3728  0.0333  0.4087  3.3024
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9356     0.1581    5.92 1.4e-08 ***
## LastAuthorFemale1  0.2389     0.1307    1.83  0.069 .
## Year1997        -0.1846     0.2989   -0.62  0.538
## Year1998         0.3337     0.3346    1.00  0.320
## Year1999         0.1112     0.2392    0.46  0.643
## Year2000        -0.0641     0.2291   -0.28  0.780
## Year2001        -0.1896     0.1934   -0.98  0.328
## Year2002         0.0441     0.2048    0.22  0.830
## Year2003         0.0435     0.1930    0.23  0.822
## Year2004        -0.0997     0.2143   -0.47  0.642
## Year2005        -0.0337     0.2346   -0.14  0.886
## Year2006        -0.0614     0.2018   -0.30  0.761
```

```

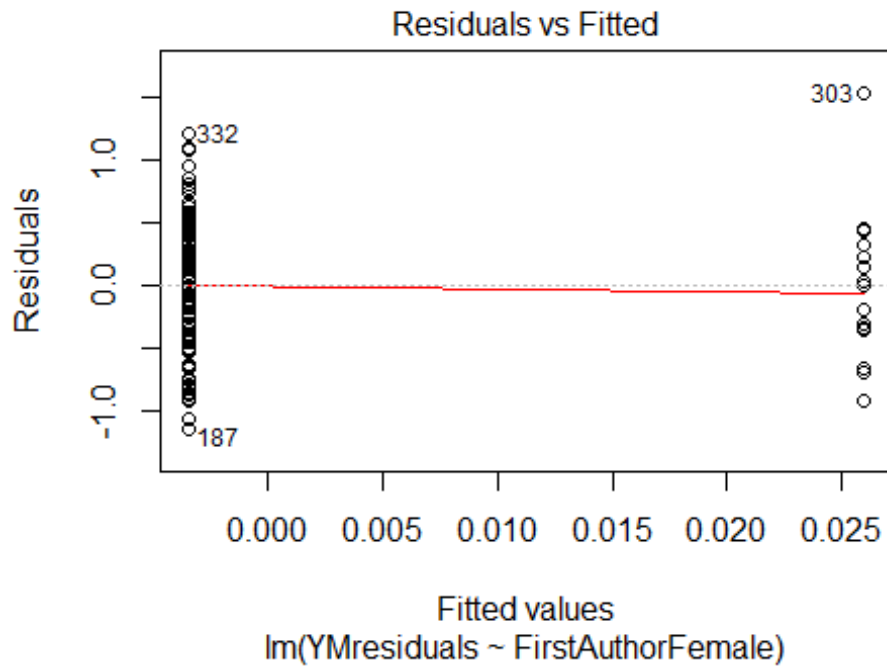
## Year2007          0.0364      0.2487      0.15      0.884
## Year2008         -0.0307      0.2238     -0.14      0.891
## Year2009          0.0456      0.2082      0.22      0.827
## Year2010          0.4352      0.2317      1.88      0.062 .
## Year2011          0.1229      0.1867      0.66      0.511
## Year2012         -0.1582      0.1855     -0.85      0.395
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.0896, Adjusted R-squared:  0.0111
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.00047);
## 20 weights are ~= 1. The remaining 194 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.398  0.876  0.953  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      4.65e-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: 215"
## [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
##    9   15   10    8   10   17   15   10   11   11   11    8   16   15   13
## 2011 2012
##    20   17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6    6    7    5   10    6    8    9    9    9    7    9   12    9

```

```
## 2011 2012
## 13 15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 4 6 7 3 9 6 4 7 8 9 6 8 12 8
## 2011 2012
## 9 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 = 8.5, df = 16, p-value = 0.9
```



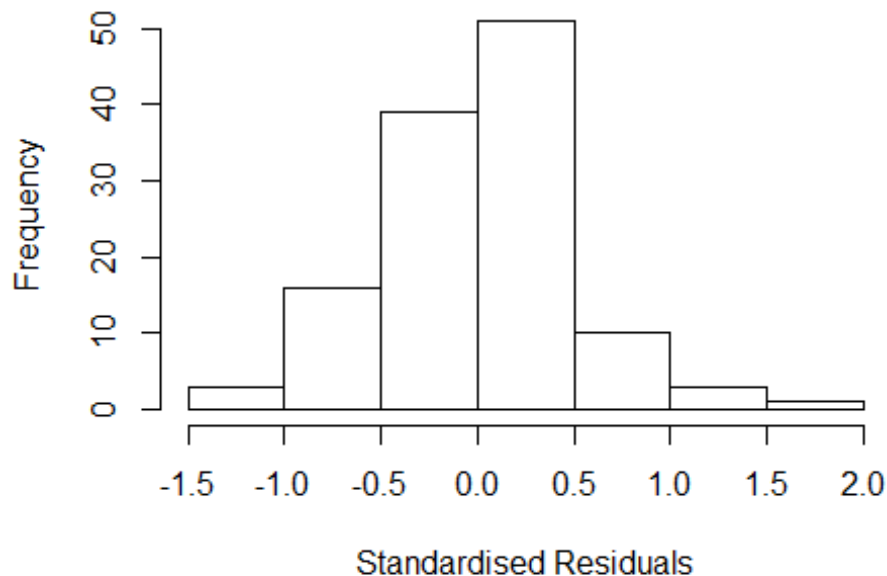
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.47, df = 1, p-value = 0.5
```

```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 1.36079000017438"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF^(1/(2*Df))
## FirstAuthorFemale	2.791	1	1.671
## LastAuthorFemale	5.104	1	2.259
## UniqueAuthors	66.286	4	1.689
## Year	92.560	16	1.152

Residuals from first and last author and team size



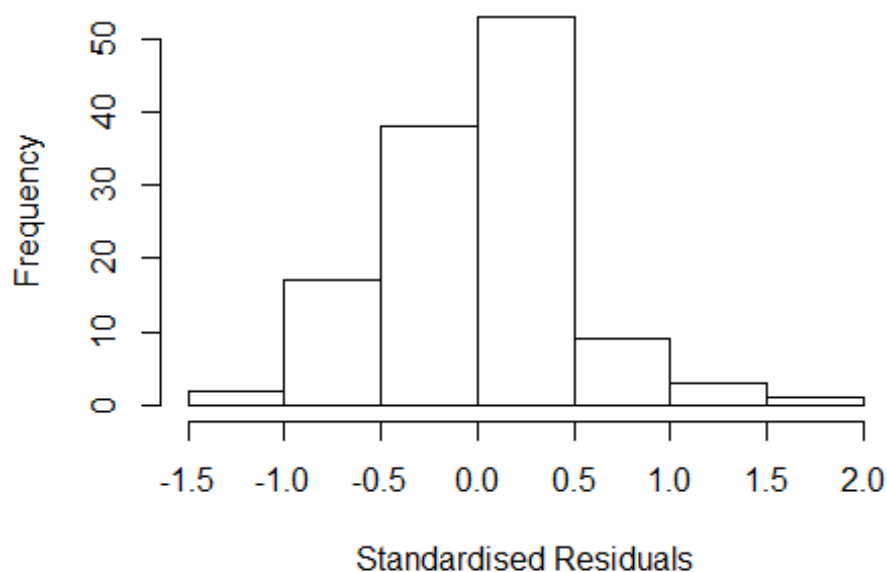
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1764 -0.2633 0.0311 0.2580 1.5554
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8591 0.2007 4.28 4.3e-05 ***
## FirstAuthorFemale1 0.1376 0.1758 0.78 0.436
## LastAuthorFemale1 -0.2486 0.1657 -1.50 0.137
## UniqueAuthors2 0.0758 0.1026 0.74 0.462
## UniqueAuthors3 0.2576 0.1486 1.73 0.086 .
## UniqueAuthors4 0.1453 0.1570 0.93 0.357
## UniqueAuthors5 0.3875 0.1875 2.07 0.041 *
## Year1997 0.1500 0.2487 0.60 0.548
## Year1998 0.1085 0.2220 0.49 0.626
## Year1999 0.2001 0.3113 0.64 0.522
```

```

## Year2000          0.6922      0.2834      2.44      0.016 *
## Year2001          0.0771      0.2931      0.26      0.793
## Year2002         -0.1438      0.2705     -0.53      0.596
## Year2003          0.1826      0.5045      0.36      0.718
## Year2004         -0.1066      0.3878     -0.27      0.784
## Year2005         -0.3445      0.2423     -1.42      0.158
## Year2006          0.2416      0.2478      0.97      0.332
## Year2007          0.3064      0.2031      1.51      0.135
## Year2008         -0.0662      0.2066     -0.32      0.749
## Year2009         -0.1492      0.2272     -0.66      0.513
## Year2010          0.2903      0.2797      1.04      0.302
## Year2011          0.0671      0.2521      0.27      0.791
## Year2012         -0.1817      0.2523     -0.72      0.473
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.468
## Multiple R-squared:  0.198, Adjusted R-squared:  0.0213
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 112 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.247  0.875   0.966   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      8.13e-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 2.391 1      1.546
## LastAuthorFemale  2.253 1      1.501
## Year              2.961 16      1.035

```

Residuals from first and last author



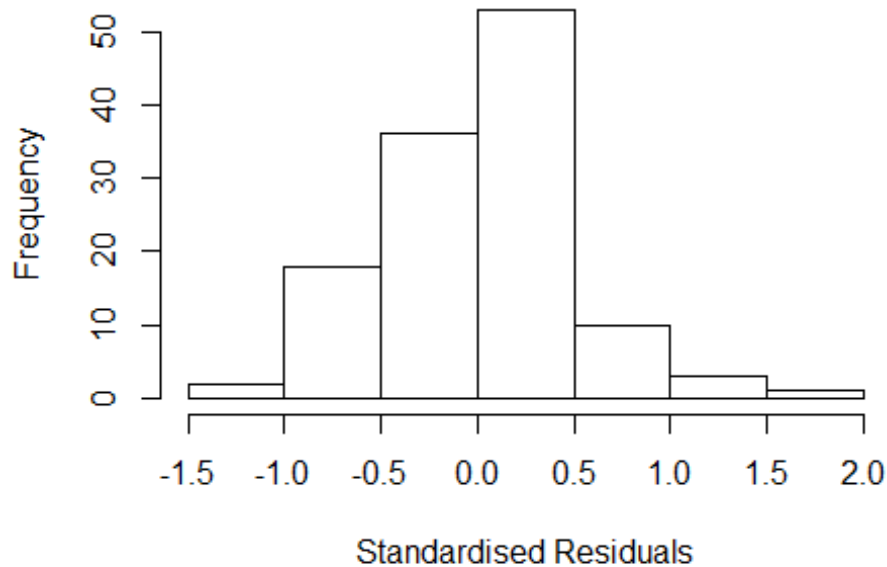
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2019 -0.2892 0.0337 0.2845 1.6007
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8970 0.1879 4.77 5.9e-06 ***
## FirstAuthorFemale1 0.1366 0.1605 0.85 0.396
## LastAuthorFemale1 -0.2735 0.1533 -1.78 0.077 .
## Year1997 0.1726 0.2247 0.77 0.444
## Year1998 0.1440 0.2193 0.66 0.513
## Year1999 0.1787 0.3151 0.57 0.572
## Year2000 0.6994 0.2690 2.60 0.011 *
## Year2001 0.1320 0.2739 0.48 0.631
## Year2002 -0.1327 0.2672 -0.50 0.621
## Year2003 0.2260 0.4568 0.49 0.622
## Year2004 -0.0532 0.3660 -0.15 0.885
## Year2005 -0.3222 0.2557 -1.26 0.210
```

```

## Year2006          0.3049      0.2461      1.24      0.218
## Year2007          0.3427      0.2033      1.69      0.095 .
## Year2008         -0.0459      0.2140     -0.21      0.831
## Year2009         -0.0766      0.2222     -0.34      0.731
## Year2010          0.3506      0.2823      1.24      0.217
## Year2011          0.0608      0.2585      0.23      0.815
## Year2012         -0.1484      0.2481     -0.60      0.551
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.465
## Multiple R-squared:  0.173, Adjusted R-squared:  0.0296
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 113 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.211  0.872  0.958  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      8.13e-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 1.851 1      1.361
## Year              1.851 16      1.019

```

Residuals from first author



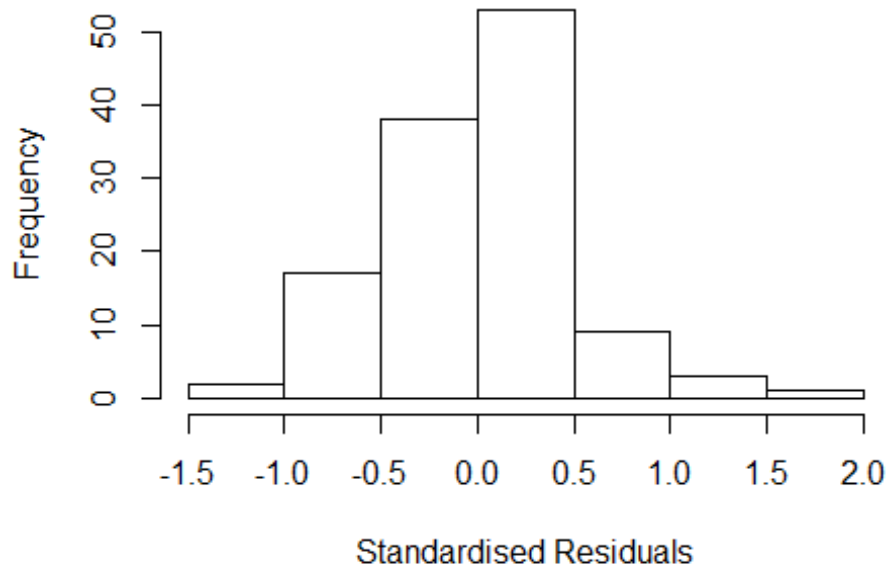
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1967 -0.2820 0.0653 0.2885 1.7136
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9176 0.1919 4.78 5.7e-06 ***
## FirstAuthorFemale1 0.0276 0.1339 0.21 0.837
## Year1997 0.0993 0.2646 0.38 0.708
## Year1998 0.0881 0.2375 0.37 0.711
## Year1999 0.1754 0.3249 0.54 0.590
## Year2000 0.5228 0.2931 1.78 0.077 .
## Year2001 0.1002 0.2706 0.37 0.712
## Year2002 -0.1545 0.2690 -0.57 0.567
## Year2003 0.2076 0.4465 0.46 0.643
## Year2004 -0.1111 0.3319 -0.33 0.738
## Year2005 -0.3411 0.2581 -1.32 0.189
## Year2006 0.2790 0.2499 1.12 0.267
```

```

## Year2007          0.2993      0.2209      1.35      0.178
## Year2008         -0.0747      0.2213     -0.34      0.736
## Year2009         -0.1302      0.2243     -0.58      0.563
## Year2010          0.3014      0.2740      1.10      0.274
## Year2011          0.0362      0.2556      0.14      0.888
## Year2012         -0.1686      0.2478     -0.68      0.498
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.483
## Multiple R-squared:  0.151, Adjusted R-squared:  0.013
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 114 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.182  0.882  0.958  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      8.13e-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 1.855 1          1.362
## Year            1.855 16          1.020

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2043 -0.2945 0.0577 0.2890 1.7586
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.93129 0.18667 4.99 2.4e-06 ***
## LastAuthorFemale1 -0.21367 0.14379 -1.49 0.140
## Year1997 0.12621 0.22840 0.55 0.582
## Year1998 0.10143 0.21945 0.46 0.645
## Year1999 0.17311 0.33320 0.52 0.604
## Year2000 0.66938 0.26302 2.54 0.012 *
## Year2001 0.12278 0.27016 0.45 0.650
## Year2002 -0.16649 0.26679 -0.62 0.534
## Year2003 0.19057 0.46240 0.41 0.681
## Year2004 -0.09193 0.36617 -0.25 0.802
## Year2005 -0.35741 0.25509 -1.40 0.164
## Year2006 0.27303 0.24519 1.11 0.268
```

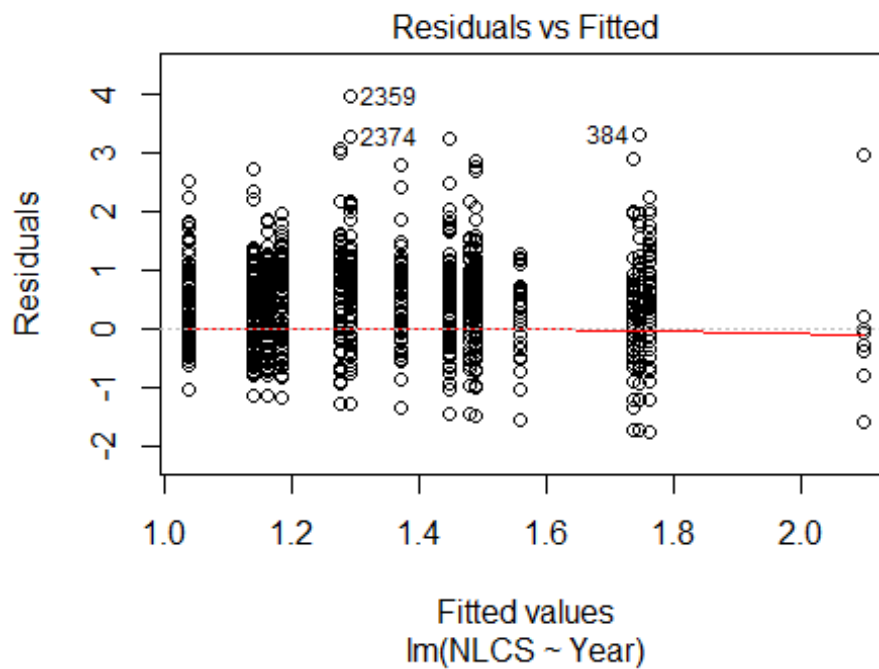


```

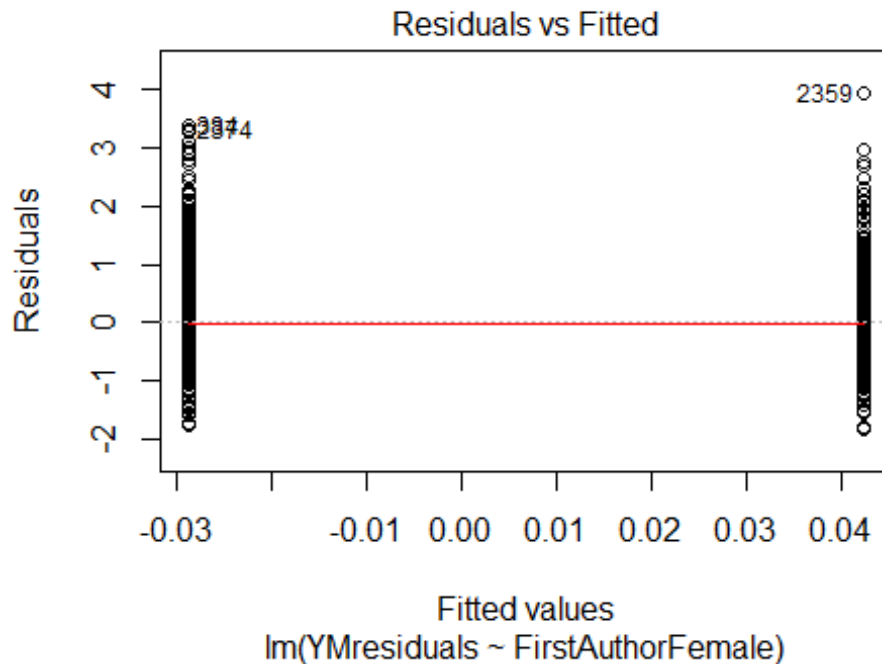
## Year2007      0.32263      0.21089      1.53      0.129
## Year2008     -0.04678      0.21570     -0.22      0.829
## Year2009     -0.09517      0.22686     -0.42      0.676
## Year2010      0.32861      0.28894      1.14      0.258
## Year2011      0.00506      0.23978      0.02      0.983
## Year2012     -0.15954      0.25658     -0.62      0.535
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.457
## Multiple R-squared:  0.171, Adjusted R-squared:  0.0364
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 115 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.105  0.878  0.958  0.896  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      8.13e-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: 123"
## [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 2008 2009 2010
## 131 137 120 110 20 162 187 164 163 171 206 230 243 284 271
## 2011 2012
## 326 286
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 74 76 60 54 8 58 142 132 143 149 162 179 198 229 200
## 2011 2012

```

```
## 250 231
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 72 71 53 48 7 55 128 120 129 132 142 163 179 202 187
## 2011 2012
## 232 214
## [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
```

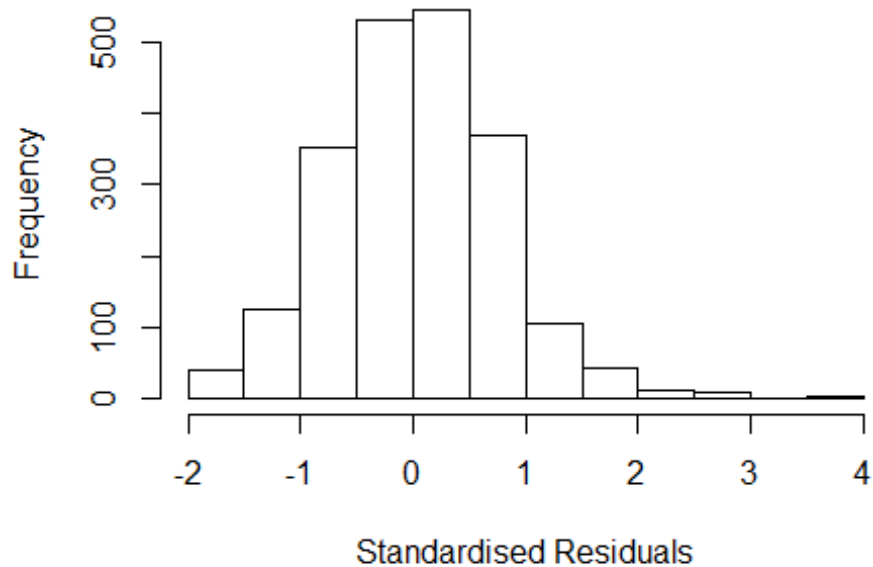


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.5, df = 1, p-value = 0.1
```



```
## [1] "Female first author team size 2018 geometric mean: 3.86995125367377"
## [1] "Male first author team size 2018 geometric mean: 3.46797798431986"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.34927799819773"
## [1] "Male last author team size 2018 geometric mean: 3.87492307732757"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1100, 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  1.157 1      1.076
## UniqueAuthors     1.197 4      1.023
## Year              1.251 16      1.007
```

Residuals from first and last author and team size



```
## [1] "List of 10 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 84      0030604595 3.711 1996      2700      1      2.519
## 283     0031034474 4.645 1997      2700      1      2.613
## 384     2642654221 5.073 1998      2700      1      3.935
## 1034    0242659412 4.230 2003      2700      1      2.599
## 1307    3242722324 4.336 2004      2700      1      2.750
## 1719    33745652285 4.159 2006      2700      1      2.526
## 2020    34249067923 4.334 2007      2700      1      2.702
## 2063    33847767624 4.263 2007      2700      1      2.616
## 2359    43049097195 5.256 2008      2700      1      3.773
## 2374    39149101470 4.570 2008      2700      1      2.933
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9455 -0.4905  0.0108  0.5056  3.9348
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1922     0.1371   8.70 < 2e-16 ***
```

```

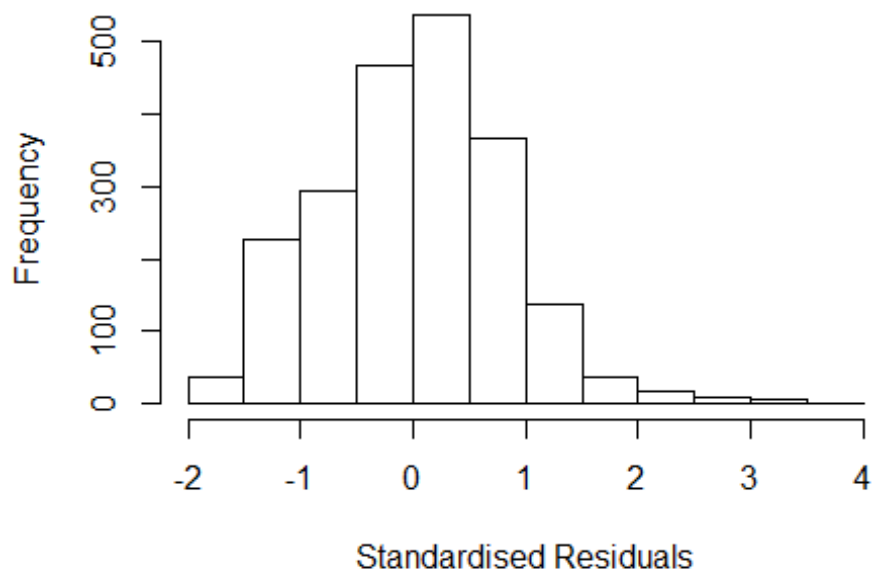
## FirstAuthorFemale1  0.0608      0.0346      1.75      0.0794 .
## LastAuthorFemale1  -0.0461      0.0367     -1.25      0.2098
## UniqueAuthors2     0.5616      0.0491     11.45 < 2e-16 ***
## UniqueAuthors3     0.6926      0.0504     13.74 < 2e-16 ***
## UniqueAuthors4     0.7592      0.0523     14.51 < 2e-16 ***
## UniqueAuthors5     0.9071      0.0552     16.43 < 2e-16 ***
## Year1997            -0.0673      0.1639     -0.41      0.6816
## Year1998            -0.0540      0.1989     -0.27      0.7860
## Year1999            -0.1728      0.1603     -1.08      0.2813
## Year2000             0.2381      0.2856      0.83      0.4045
## Year2001            -0.2171      0.1731     -1.25      0.2098
## Year2002            -0.3175      0.1466     -2.17      0.0304 *
## Year2003            -0.2079      0.1501     -1.39      0.1662
## Year2004            -0.2985      0.1466     -2.04      0.0418 *
## Year2005            -0.3525      0.1463     -2.41      0.0160 *
## Year2006            -0.4810      0.1464     -3.29      0.0010 **
## Year2007            -0.4670      0.1493     -3.13      0.0018 **
## Year2008            -0.4624      0.1439     -3.21      0.0013 **
## Year2009            -0.5684      0.1423     -4.00      6.7e-05 ***
## Year2010            -0.6526      0.1408     -4.64      3.8e-06 ***
## Year2011            -0.7163      0.1382     -5.18      2.4e-07 ***
## Year2012            -0.7072      0.1382     -5.12      3.4e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.714
## Multiple R-squared:  0.219, Adjusted R-squared:  0.211
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 2 observations c(175,1272) are outliers with |weight| = 0 ( < 4.7e-05);
## 180 weights are ~ = 1. The remaining 1952 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0531 0.8750 0.9490 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           4.69e-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 1.088 1          1.043
## LastAuthorFemale  1.078 1          1.038
## Year              1.050 16          1.002
```

Residuals from first and last author



```
## [1] "List of 13 outliers with residuals above 2.5"
```

```
##      ScopusId  NLCS Year OneField Fields residuals
## 283   0031034474 4.645 1997    2700      1    2.929
## 384   2642654221 5.073 1998    2700      1    3.387
## 1034  0242659412 4.230 2003    2700      1    2.846
## 1223  6444225735 4.176 2004    2700      1    2.712
## 1307  3242722324 4.336 2004    2700      1    2.872
## 1515  19744379509 3.929 2005    2700      1    2.669
## 1719  33745652285 4.159 2006    2700      1    2.846
## 2020  34249067923 4.334 2007    2700      1    3.104
## 2063  33847767624 4.263 2007    2700      1    3.031
## 2359  43049097195 5.256 2008    2700      1    3.903
## 2374  39149101470 4.570 2008    2700      1    3.337
## 3304  79957951981 3.545 2011    2700      1    2.551
## 3762  84861516992 3.855 2012    2700      1    2.629
```

```
##
```

```
## Call:
```

```
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
```

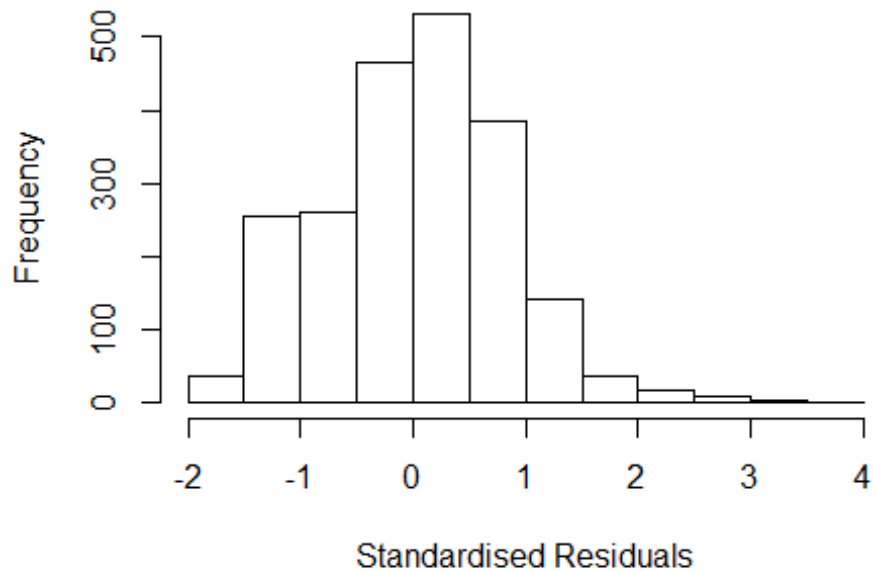
```

## \--> method = "MM"
## Residuals:
##   Min      1Q  Median      3Q      Max
## -1.894 -0.542  0.053  0.562  3.903
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7744     0.1389   12.77 < 2e-16 ***
## FirstAuthorFemale1  0.1197     0.0371    3.22 0.00128 **
## LastAuthorFemale1 -0.1180     0.0409   -2.89 0.00395 **
## Year1997          -0.0579     0.1666   -0.35 0.72805
## Year1998          -0.0885     0.1974   -0.45 0.65396
## Year1999          -0.2598     0.1690   -1.54 0.12444
## Year2000          -0.0709     0.2544   -0.28 0.78061
## Year2001          -0.3051     0.1789   -1.71 0.08829 .
## Year2002          -0.3128     0.1564   -2.00 0.04566 *
## Year2003          -0.2728     0.1627   -1.68 0.09379 .
## Year2004          -0.3099     0.1561   -1.98 0.04729 *
## Year2005          -0.3963     0.1540   -2.57 0.01011 *
## Year2006          -0.4633     0.1555   -2.98 0.00291 **
## Year2007          -0.5444     0.1567   -3.47 0.00052 ***
## Year2008          -0.5412     0.1543   -3.51 0.00046 ***
## Year2009          -0.6403     0.1489   -4.30 1.8e-05 ***
## Year2010          -0.6626     0.1485   -4.46 8.6e-06 ***
## Year2011          -0.7823     0.1462   -5.35 9.7e-08 ***
## Year2012          -0.6677     0.1458   -4.58 4.9e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.801
## Multiple R-squared:  0.0697, Adjusted R-squared:  0.0617
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 1272 is an outlier with |weight| = 0 ( < 4.7e-05);
## 170 weights are ~ = 1. The remaining 1963 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0347 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           4.69e-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 1.029 1          1.014
## Year              1.029 16          1.001
```

Residuals from first author



```
## [1] "List of 13 outliers with residuals above 2.5"
##          ScopusId NLCS Year OneField Fields residuals
## 283      0031034474 4.645 1997      2700      1      2.929
## 384      2642654221 5.073 1998      2700      1      3.387
## 1034     0242659412 4.230 2003      2700      1      2.846
## 1223     6444225735 4.176 2004      2700      1      2.712
## 1307     3242722324 4.336 2004      2700      1      2.872
## 1515     19744379509 3.929 2005      2700      1      2.669
## 1719     33745652285 4.159 2006      2700      1      2.846
## 2020     34249067923 4.334 2007      2700      1      3.104
## 2063     33847767624 4.263 2007      2700      1      3.031
## 2359     43049097195 5.256 2008      2700      1      3.903
## 2374     39149101470 4.570 2008      2700      1      3.337
## 3304     79957951981 3.545 2011      2700      1      2.551
## 3762     84861516992 3.855 2012      2700      1      2.629
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
```

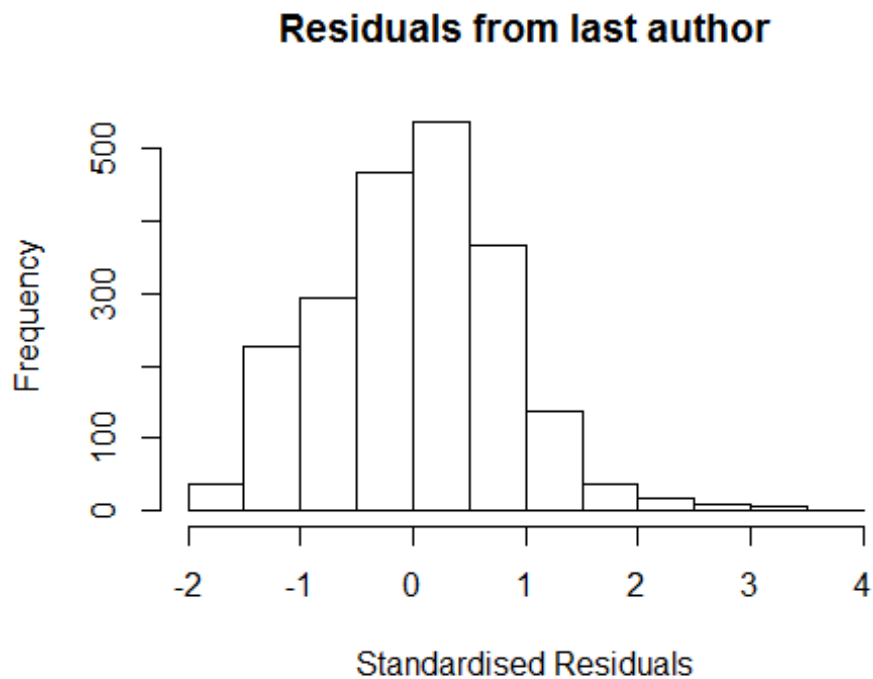


```

AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.8338 -0.5433  0.0422  0.5525  3.9635
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7466      0.1389   12.58 < 2e-16 ***
## FirstAuthorFemale1  0.0873      0.0365    2.39  0.01673 *
## Year1997         -0.0536      0.1683   -0.32  0.74991
## Year1998         -0.0837      0.1971   -0.42  0.67104
## Year1999         -0.2435      0.1681   -1.45  0.14772
## Year2000         -0.0820      0.2559   -0.32  0.74871
## Year2001         -0.2987      0.1781   -1.68  0.09362 .
## Year2002         -0.3075      0.1566   -1.96  0.04968 *
## Year2003         -0.2697      0.1629   -1.66  0.09800 .
## Year2004         -0.3030      0.1565   -1.94  0.05300 .
## Year2005         -0.3835      0.1541   -2.49  0.01293 *
## Year2006         -0.4593      0.1556   -2.95  0.00320 **
## Year2007         -0.5377      0.1575   -3.41  0.00065 ***
## Year2008         -0.5414      0.1552   -3.49  0.00050 ***
## Year2009         -0.6342      0.1495   -4.24  2.3e-05 ***
## Year2010         -0.6524      0.1492   -4.37  1.3e-05 ***
## Year2011         -0.7745      0.1467   -5.28  1.4e-07 ***
## Year2012         -0.6621      0.1462   -4.53  6.3e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.801
## Multiple R-squared:  0.0655, Adjusted R-squared:  0.058
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 1272 is an outlier with |weight| = 0 ( < 4.7e-05);
## 164 weights are ~= 1. The remaining 1969 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.0303  0.8630  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          4.69e-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 1.017 1          1.009
## Year             1.017 16          1.001
```



```
## [1] "List of 13 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 283   0031034474 4.645 1997   2700      1    2.929
## 384   2642654221 5.073 1998   2700      1    3.387
## 1034  0242659412 4.230 2003   2700      1    2.846
## 1223  6444225735 4.176 2004   2700      1    2.712
## 1307  3242722324 4.336 2004   2700      1    2.872
## 1515  19744379509 3.929 2005   2700      1    2.669
## 1719  33745652285 4.159 2006   2700      1    2.846
## 2020  34249067923 4.334 2007   2700      1    3.104
## 2063  33847767624 4.263 2007   2700      1    3.031
## 2359  43049097195 5.256 2008   2700      1    3.903
## 2374  39149101470 4.570 2008   2700      1    3.337
## 3304  79957951981 3.545 2011   2700      1    2.551
## 3762  84861516992 3.855 2012   2700      1    2.629
##
## Call:
```

```

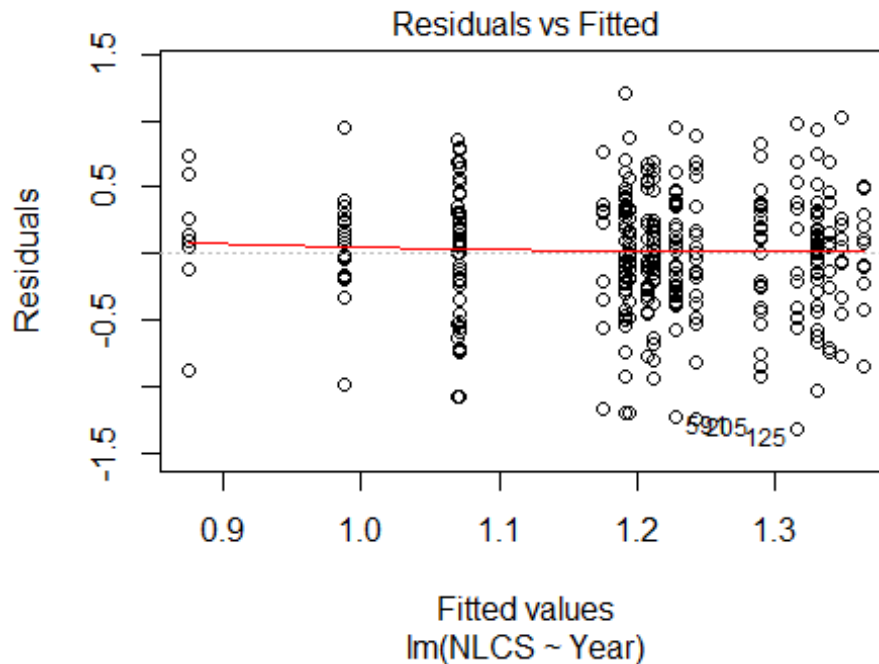
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.787 -0.542  0.049  0.564  3.985
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.7874     0.1395   12.81 < 2e-16 ***
## LastAuthorFemale1 -0.0799     0.0393   -2.03  0.04239 *
## Year1997         -0.0417     0.1683   -0.25  0.80431
## Year1998         -0.0691     0.1968   -0.35  0.72554
## Year1999         -0.2413     0.1676   -1.44  0.15010
## Year2000         -0.0469     0.2546   -0.18  0.85396
## Year2001         -0.2803     0.1799   -1.56  0.11933
## Year2002         -0.2905     0.1565   -1.86  0.06352 .
## Year2003         -0.2648     0.1637   -1.62  0.10603
## Year2004         -0.2806     0.1568   -1.79  0.07367 .
## Year2005         -0.3732     0.1547   -2.41  0.01592 *
## Year2006         -0.4403     0.1558   -2.83  0.00476 **
## Year2007         -0.5189     0.1572   -3.30  0.00098 ***
## Year2008         -0.5160     0.1551   -3.33  0.00090 ***
## Year2009         -0.6049     0.1493   -4.05  5.3e-05 ***
## Year2010         -0.6330     0.1494   -4.24  2.4e-05 ***
## Year2011         -0.7574     0.1469   -5.16  2.8e-07 ***
## Year2012         -0.6388     0.1465   -4.36  1.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.804
## Multiple R-squared:  0.0653, Adjusted R-squared:  0.0578
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 1272 is an outlier with |weight| = 0 ( < 4.7e-05);
## 184 weights are ~ = 1. The remaining 1949 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.0426  0.8590  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          4.69e-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: 2134"
## [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
##    26    22    18    21    20    37    28    30    21    28    35    46    33    43    50
## 2011 2012
##    40    58
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    10    10    11     8    12    10    21    21     9    20    23    33    25    31    42
## 2011 2012
##    32    43
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     9     8     9     7     9     8    20    17     8    18    19    32    23    28    38
## 2011 2012
##    25    38
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.6

```



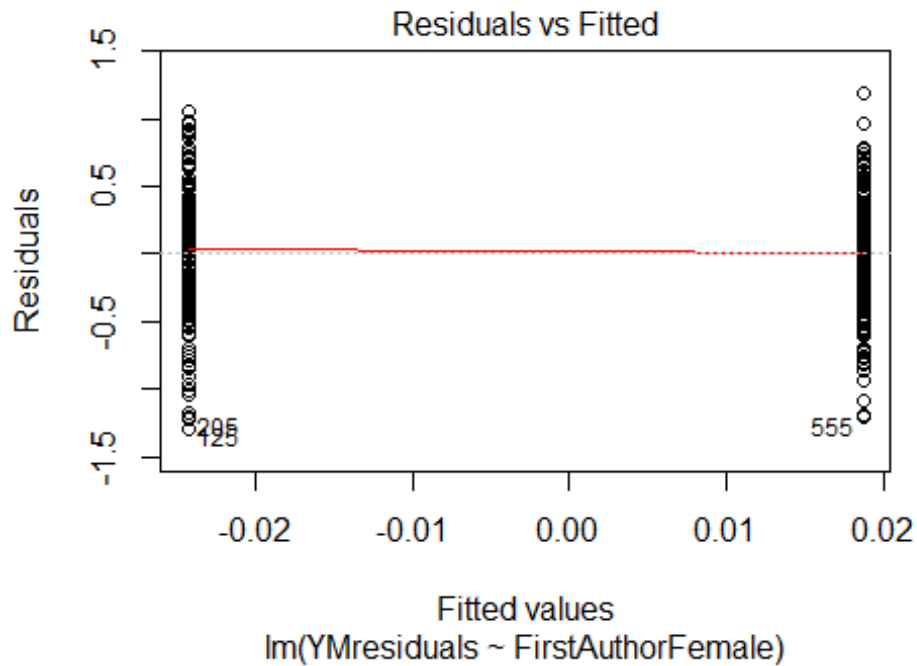
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8.6, df = 1, p-value = 0.003

## [1] "Female first author team size 2018 geometric mean: 4.6331326865723"
## [1] "Male first author team size 2018 geometric mean: 4.0399111803486"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

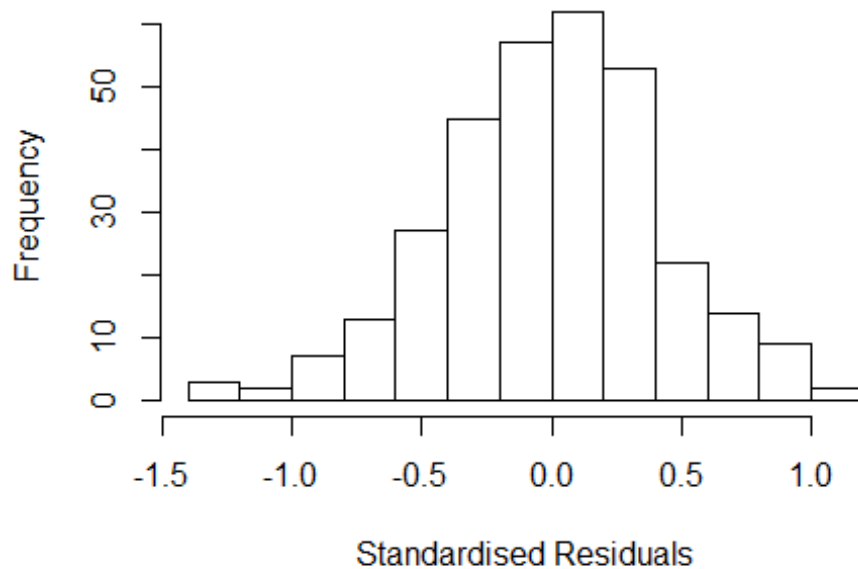
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.28225473667665"
## [1] "Male last author team size 2018 geometric mean: 4.47068643765029"

## 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 = 0.5
## 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 1.384 1          1.176
## LastAuthorFemale  1.226 1          1.107
## UniqueAuthors    3.691 4          1.177
## Year              4.920 16         1.051
```

Residuals from first and last author and team size



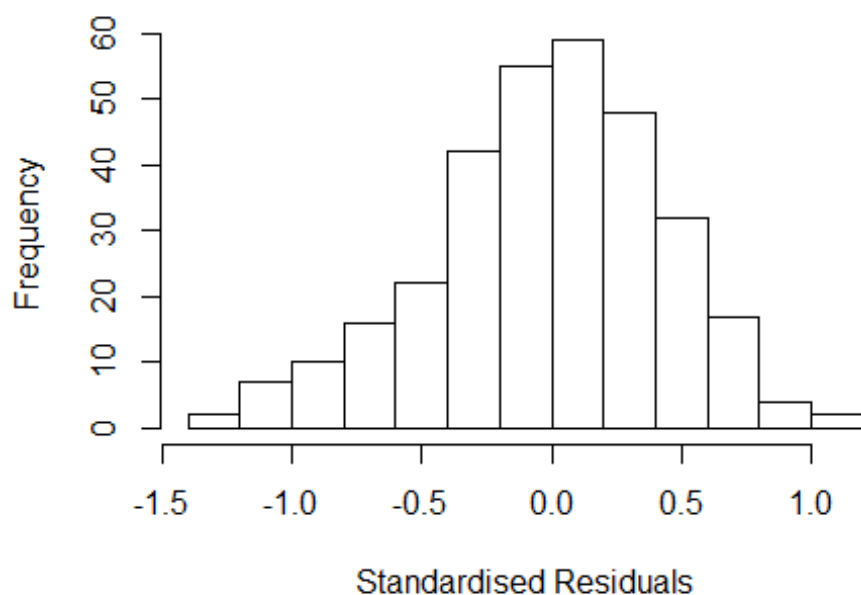
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36165 -0.27626  0.00701  0.26424  1.13239
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.85748    0.17533   4.89 1.7e-06 ***
## FirstAuthorFemale1 0.03094    0.05500   0.56 0.57423
## LastAuthorFemale1 0.05134    0.05104   1.01 0.31535
## UniqueAuthors2    0.35493    0.12309   2.88 0.00422 **
## UniqueAuthors3    0.41786    0.10641   3.93 0.00011 ***
## UniqueAuthors4    0.46237    0.11282   4.10 5.4e-05 ***
## UniqueAuthors5    0.38656    0.10722   3.61 0.00037 ***
## Year1997          0.07491    0.17170   0.44 0.66295
## Year1998          0.16096    0.19085   0.84 0.39969
## Year1999         -0.04047    0.25651  -0.16 0.87476
```

```

## Year2000          0.18640      0.27733      0.67  0.50202
## Year2001          -0.14470      0.23446     -0.62  0.53762
## Year2002           0.10046      0.18307      0.55  0.58359
## Year2003          -0.11205      0.18615     -0.60  0.54769
## Year2004          -0.33113      0.24532     -1.35  0.17812
## Year2005          -0.25326      0.17347     -1.46  0.14538
## Year2006          -0.05287      0.17005     -0.31  0.75608
## Year2007          -0.06806      0.16839     -0.40  0.68636
## Year2008          -0.00822      0.18597     -0.04  0.96479
## Year2009           0.05064      0.17165      0.30  0.76819
## Year2010          -0.09301      0.16956     -0.55  0.58373
## Year2011           0.02849      0.17958      0.16  0.87405
## Year2012          -0.15180      0.17466     -0.87  0.38548
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.169, Adjusted R-squared:  0.106
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 291 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.230  0.862  0.951  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      3.16e-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 1.436 1      1.198
## LastAuthorFemale  1.288 1      1.135
## Year              1.782 16      1.018

```


Residuals from first and last author



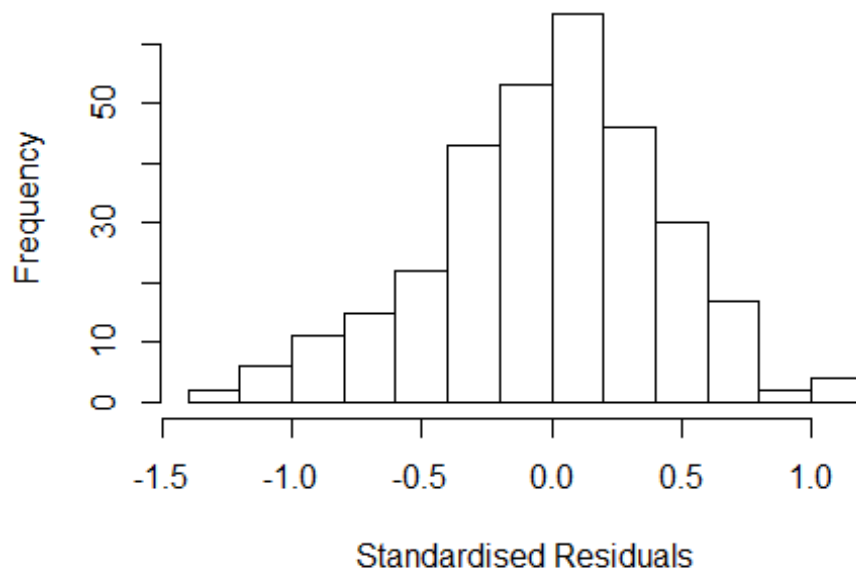
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2714 -0.2689 0.0142 0.2802 1.1256
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2585 0.1540 8.17 9e-15 ***
## FirstAuthorFemale1 0.0710 0.0581 1.22 0.222
## LastAuthorFemale1 0.0521 0.0547 0.95 0.341
## Year1997 -0.0785 0.2125 -0.37 0.712
## Year1998 0.0225 0.2058 0.11 0.913
## Year1999 -0.3216 0.3497 -0.92 0.358
## Year2000 -0.0885 0.2857 -0.31 0.757
## Year2001 -0.2289 0.2443 -0.94 0.350
## Year2002 -0.0813 0.1868 -0.44 0.664
## Year2003 -0.1480 0.1899 -0.78 0.436
## Year2004 -0.4281 0.2884 -1.48 0.139
## Year2005 -0.3398 0.1832 -1.85 0.065 .
```

```

## Year2006          -0.1641      0.1781   -0.92    0.358
## Year2007          -0.0994      0.1767   -0.56    0.574
## Year2008          -0.0574      0.1977   -0.29    0.772
## Year2009          -0.0123      0.1824   -0.07    0.946
## Year2010          -0.1103      0.1794   -0.61    0.539
## Year2011          -0.0360      0.1863   -0.19    0.847
## Year2012          -0.2407      0.1815   -1.33    0.186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.0667, Adjusted R-squared:  0.0101
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 292 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.344  0.868  0.953  0.899  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      3.16e-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 1.491 1          1.221
## Year              1.491 16          1.013

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.25850 -0.26324 0.00887 0.27627 1.13850
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2648 0.1525 8.29 3.8e-15 ***
## FirstAuthorFemale1 0.0875 0.0601 1.46 0.147
## Year1997 -0.0712 0.2072 -0.34 0.732
## Year1998 0.0359 0.2087 0.17 0.863
## Year1999 -0.3176 0.3413 -0.93 0.353
## Year2000 -0.0989 0.2843 -0.35 0.728
## Year2001 -0.2016 0.2432 -0.83 0.408
## Year2002 -0.0703 0.1868 -0.38 0.707
## Year2003 -0.1352 0.1884 -0.72 0.474
## Year2004 -0.3980 0.2816 -1.41 0.159
## Year2005 -0.3309 0.1828 -1.81 0.071 .
## Year2006 -0.1539 0.1758 -0.88 0.382
```

```

## Year2007          -0.0874      0.1745   -0.50    0.617
## Year2008          -0.0447      0.1949   -0.23    0.819
## Year2009           0.0016      0.1800    0.01    0.993
## Year2010          -0.0938      0.1776   -0.53    0.598
## Year2011          -0.0328      0.1840   -0.18    0.859
## Year2012          -0.2348      0.1801   -1.30    0.193
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.422
## Multiple R-squared:  0.0633, Adjusted R-squared:  0.00985
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 285 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.353  0.865  0.955  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      3.16e-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 1.364 1      1.168
## Year      1.364 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
## -1.2625 -0.2827  0.0114  0.2772  1.1460

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.26787    0.15154    8.37 2.3e-15 ***
## LastAuthorFemale1 0.07366    0.05743    1.28  0.201
## Year1997       -0.05719    0.20481   -0.28  0.780
## Year1998        0.04773    0.19481    0.24  0.807
## Year1999       -0.31054    0.35270   -0.88  0.379
## Year2000       -0.09083    0.29505   -0.31  0.758
## Year2001       -0.21608    0.24463   -0.88  0.378
## Year2002       -0.07159    0.18492   -0.39  0.699
## Year2003       -0.13895    0.18589   -0.75  0.455
## Year2004       -0.40510    0.28942   -1.40  0.163
## Year2005       -0.32125    0.17575   -1.83  0.069 .
## Year2006       -0.13647    0.16847   -0.81  0.419
## Year2007       -0.07903    0.16841   -0.47  0.639
## Year2008       -0.02625    0.18847   -0.14  0.889
## Year2009        0.01288    0.17302    0.07  0.941
## Year2010       -0.09057    0.17083   -0.53  0.596
## Year2011       -0.00533    0.17490   -0.03  0.976
## Year2012       -0.21080    0.17117   -1.23  0.219
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.42
## Multiple R-squared:  0.061, Adjusted R-squared:  0.00745
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 291 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.345  0.859  0.954  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      3.16e-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: 316"
## [1] ""

```

```

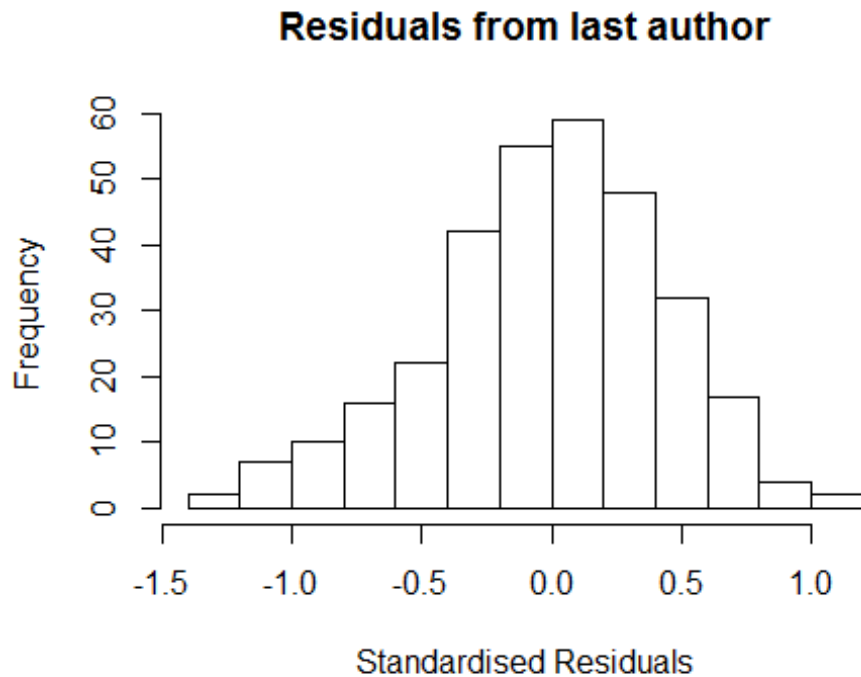
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2702"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    3    8    9   10    8   10    9   10    9   12    6    8   10
## 2011 2012
##    4   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    1    0    3    4    2    4    6    5    8    8   10    5    5    8
## 2011 2012
##    4    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    1    0    3    4    2    4    5    2    7    6    8    4    5    7
## 2011 2012
##    4    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male first author team size 2018 geometric mean: 1.44224957030741"

## 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 = 3, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.81712059283214"
## [1] "Male last author team size 2018 geometric mean: 1"

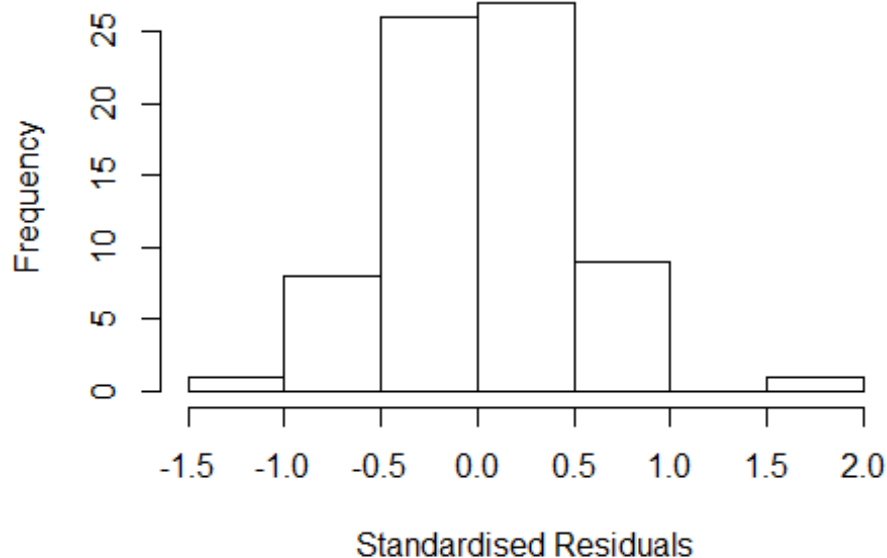
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5, 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    7.408  1      2.722
## LastAuthorFemale    19.536  1      4.420
## UniqueAuthors    1064.389  4      2.390
## Year           8772.411 15      1.353
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.2272 -0.2635 0.0196 0.2747 1.7801
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.12892 0.68072 1.66 0.10
## FirstAuthorFemale1 0.09512 0.14709 0.65 0.52
## LastAuthorFemale1 0.09162 0.15279 0.60 0.55
## UniqueAuthors2 -0.01484 0.28643 -0.05 0.96
## UniqueAuthors3 -0.20908 0.23171 -0.90 0.37
## UniqueAuthors4 -0.02672 0.25961 -0.10 0.92
## UniqueAuthors5 0.14076 0.24746 0.57 0.57
## Year1997 -0.47120 0.43475 -1.08 0.28
## Year1999 -0.22958 0.64374 -0.36 0.72
## Year2000 -0.15083 0.66922 -0.23 0.82
```

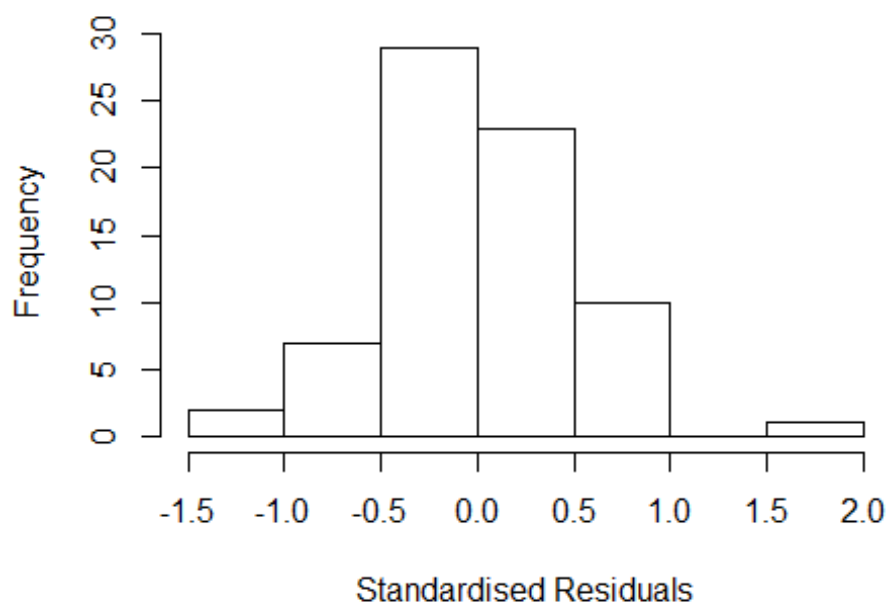


```

## Year2001      0.20374      0.59075      0.34      0.73
## Year2002     -0.30870      0.53262     -0.58      0.56
## Year2003     -0.21459      0.57286     -0.37      0.71
## Year2004     -0.20359      0.52944     -0.38      0.70
## Year2005      0.01801      0.66175      0.03      0.98
## Year2006      0.30330      0.52180      0.58      0.56
## Year2007     -0.15653      0.61827     -0.25      0.80
## Year2008      0.04853      0.53303      0.09      0.93
## Year2009     -0.15924      0.77601     -0.21      0.84
## Year2010     -0.24418      0.52983     -0.46      0.65
## Year2011     -0.00472      0.63127     -0.01      0.99
## Year2012      0.04562      0.51915      0.09      0.93
##
## Robust residual standard error: 0.493
## Multiple R-squared:  0.192, Adjusted R-squared:  -0.148
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 65 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.166  0.876  0.970   0.918   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      1.39e-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  5.674  1      2.382
## LastAuthorFemale  22.153  1      4.707
## Year      108.084 15      1.169

```

Residuals from first and last author



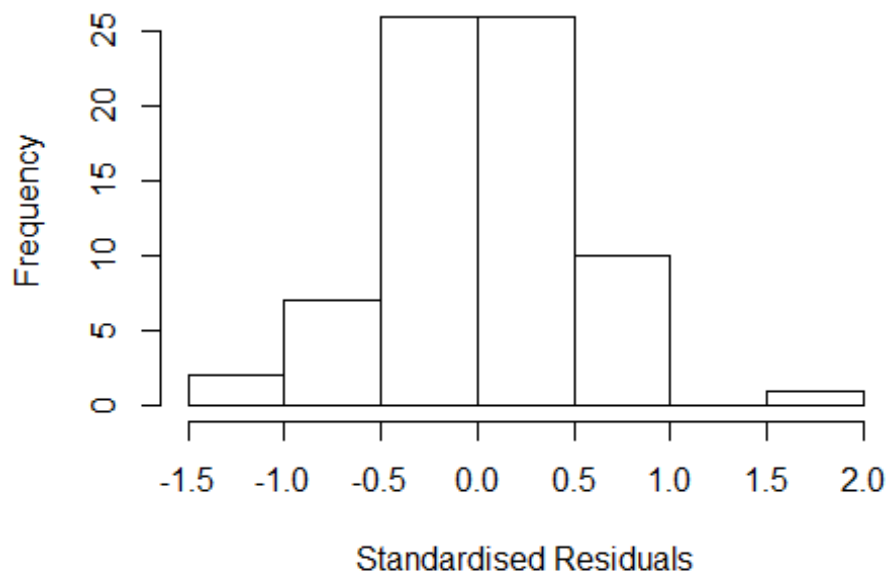
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1466 -0.2399 -0.0118 0.2548 1.8156
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09340 0.40550 2.70 0.0093 **
## FirstAuthorFemale1 0.07108 0.13499 0.53 0.6006
## LastAuthorFemale1 0.14266 0.15508 0.92 0.3617
## Year1997 -0.42648 0.38147 -1.12 0.2685
## Year1999 -0.16991 0.41125 -0.41 0.6811
## Year2000 -0.16684 0.44714 -0.37 0.7105
## Year2001 0.12123 0.37316 0.32 0.7465
## Year2002 -0.29231 0.40103 -0.73 0.4692
## Year2003 -0.19630 0.51558 -0.38 0.7049
## Year2004 -0.40414 0.35615 -1.13 0.2615
## Year2005 -0.01788 0.43251 -0.04 0.9672
## Year2006 0.32066 0.39845 0.80 0.4245
```

```

## Year2007      -0.08147    0.42367   -0.19    0.8482
## Year2008      0.08208    0.43901    0.19    0.8524
## Year2009     -0.18901    0.57638   -0.33    0.7442
## Year2010     -0.34338    0.38258   -0.90    0.3734
## Year2011     -0.00269    0.50112   -0.01    0.9957
## Year2012     -0.01074    0.38600   -0.03    0.9779
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.492
## Multiple R-squared:  0.154, Adjusted R-squared:  -0.112
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.144  0.880  0.970  0.917  0.995  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 79.33 1      8.907
## Year              79.33 15      1.157

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.26928 -0.27090 0.00422 0.29911 1.77780
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1312 0.4406 2.57 0.013 *
## FirstAuthorFemale1 0.1021 0.1300 0.79 0.436
## Year1997 -0.4953 0.4081 -1.21 0.230
## Year1999 -0.1675 0.4622 -0.36 0.718
## Year2000 -0.1735 0.4970 -0.35 0.728
## Year2001 0.1393 0.4197 0.33 0.741
## Year2002 -0.2983 0.4481 -0.67 0.508
## Year2003 -0.2181 0.5297 -0.41 0.682
## Year2004 -0.3303 0.4099 -0.81 0.424
## Year2005 0.0360 0.4622 0.08 0.938
## Year2006 0.2628 0.4298 0.61 0.543
## Year2007 -0.1021 0.4663 -0.22 0.827
```

```

## Year2008          0.0310      0.4670      0.07      0.947
## Year2009         -0.2323      0.5996     -0.39      0.700
## Year2010         -0.3449      0.4267     -0.81      0.422
## Year2011         -0.0453      0.5268     -0.09      0.932
## Year2012         -0.0251      0.4304     -0.06      0.954
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared:  0.148, Adjusted R-squared:  -0.0997
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 63 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.139  0.869  0.962  0.907  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      1.39e-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 24.66  1          4.966
## Year              24.66 15          1.113

## [1] "List of 0 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
## -1.1121 -0.2473 -0.0162  0.2639  1.7744
##

```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1346    0.4402   2.58  0.013 *
## LastAuthorFemale1 0.1613    0.1492   1.08  0.284
## Year1997         -0.3966    0.4402  -0.90  0.372
## Year1999         -0.1924    0.4590  -0.42  0.677
## Year2000         -0.1939    0.4909  -0.40  0.694
## Year2001          0.1063    0.4176   0.25  0.800
## Year2002         -0.3192    0.4373  -0.73  0.469
## Year2003         -0.2170    0.5477  -0.40  0.694
## Year2004         -0.3929    0.4076  -0.96  0.339
## Year2005         -0.0225    0.4711  -0.05  0.962
## Year2006          0.3253    0.4469   0.73  0.470
## Year2007         -0.1057    0.4606  -0.23  0.819
## Year2008          0.0648    0.4785   0.14  0.893
## Year2009         -0.2141    0.6244  -0.34  0.733
## Year2010         -0.3399    0.4360  -0.78  0.439
## Year2011         -0.0326    0.5361  -0.06  0.952
## Year2012         -0.0199    0.4308  -0.05  0.963
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.148, Adjusted R-squared:  -0.1
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 64 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.179  0.885   0.968   0.915   0.992   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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: 72"
## [1] ""
## [1] ""
## [1] "#####"

```

```

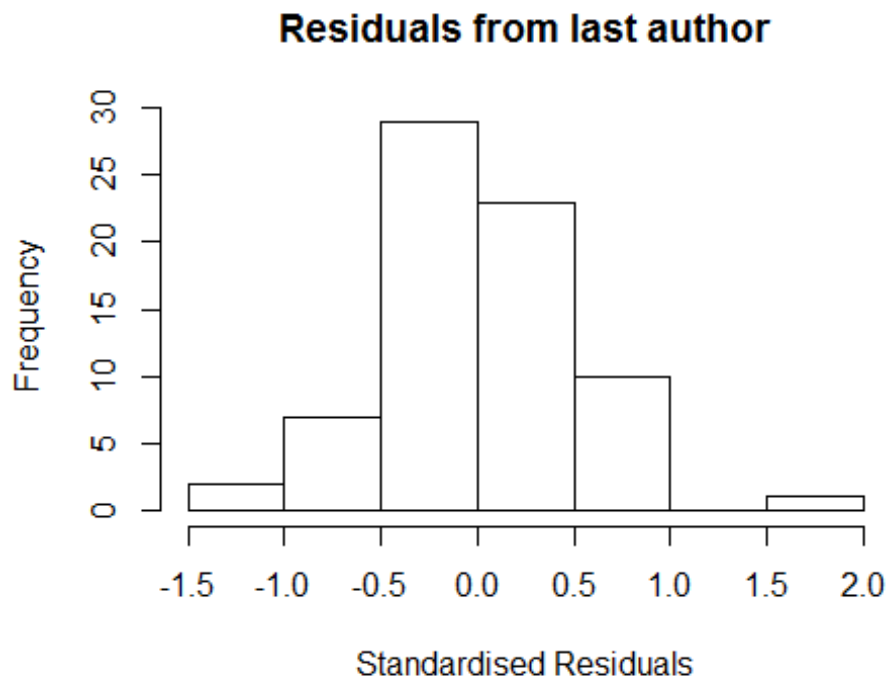
## [1] "Analysis of AJSC 2703"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 27 21 28 14 19 14 16 19 16 13 19 14 27 17
## 2011 2012
## 23 21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 6 6 2 3 2 2 5 6 3 9 8 6 18 6
## 2011 2012
## 10 12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 5 6 2 3 2 1 3 3 3 6 5 6 14 6
## 2011 2012
## 10 11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.61789292741588"
## [1] "Male first author team size 2018 geometric mean: 2.90419122184086"

## 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 = 22, p-value = 0.2
## 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: 3.39551532168461"

## 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 = 12, 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 3.738e+00  1      1.933
## LastAuthorFemale  3.055e+01  1      5.527
## UniqueAuthors    6.414e+14  4     70.940
## Year              8.226e+15 16      3.143
##
## [1] "List of 0 outliers with 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
## -1.21e+00 -3.19e-01  1.33e-15  3.08e-01  8.77e-01
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7924      0.2067   3.83 0.00029 ***
## FirstAuthorFemale1 0.1912      0.1674   1.14 0.25760
## LastAuthorFemale1 0.0726      0.1586   0.46 0.64866
## UniqueAuthors2     0.2676      0.2067   1.29 0.20007
## UniqueAuthors3     0.6036      0.1920   3.14 0.00253 **
## UniqueAuthors4     0.2419      0.2021   1.20 0.23579
## UniqueAuthors5     0.5723      0.1814   3.15 0.00245 **
## Year1997          -0.2067      0.3259  -0.63 0.52818
## Year1998           0.2982      0.2388   1.25 0.21633
## Year1999           0.3055      0.5225   0.58 0.56078
## Year2000           0.0115      0.3073   0.04 0.97027
## Year2001           0.2405      0.1855   1.30 0.19947
## Year2002          -1.0940      0.2131  -5.13 2.9e-06 ***
## Year2003           0.3454      0.4552   0.76 0.45078
## Year2004           0.0869      0.2923   0.30 0.76711
## Year2005          -0.0733      0.1033  -0.71 0.48077
## Year2006          -0.0801      0.2389  -0.34 0.73836
## Year2007           0.0786      0.2648   0.30 0.76756
## Year2008           0.5598      0.2289   2.45 0.01722 *
## Year2009           0.0419      0.1305   0.32 0.74927
## Year2010           0.1708      0.3033   0.56 0.57536
## Year2011          -0.1822      0.2509  -0.73 0.47034
## Year2012          -0.0663      0.3631  -0.18 0.85581
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.356, Adjusted R-squared:  0.135
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 74 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.508  0.908  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          1.15e-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

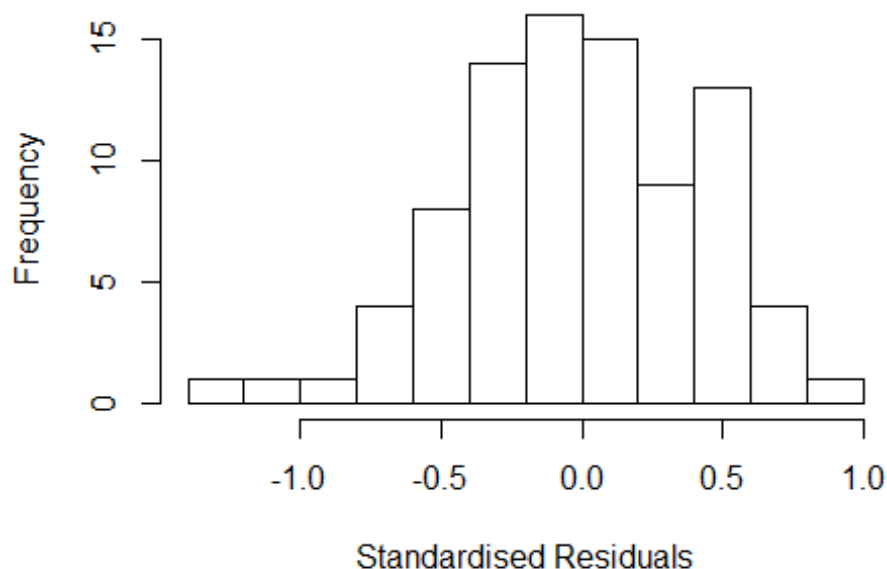
```

```
##                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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author and team size



```
##                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 = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3125 -0.3140 -0.0118  0.3220  1.1085
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0600     0.0000      Inf < 2e-16 ***
## FirstAuthorFemale1  0.1260     0.1425     0.88  0.379
## LastAuthorFemale1  0.1424     0.1575     0.90  0.369
## Year1997          -0.1025     0.2367    -0.43  0.666
## Year1998           0.3209     0.1938     1.66  0.102
## Year1999           0.3240     1.1881     0.27  0.786
## Year2000          -0.0311     0.2702    -0.12  0.909
## Year2001           0.4085     0.2942     1.39  0.170
## Year2002          -0.7580     0.0000    -Inf < 2e-16 ***
## Year2003           0.5230     0.1696     3.08  0.003 **
## Year2004           0.2735     0.1447     1.89  0.063 .
## Year2005          -0.0193     0.1363    -0.14  0.888
## Year2006           0.0230     0.2859     0.08  0.936
## Year2007           0.1108     0.2789     0.40  0.692
## Year2008           0.6222     0.0813     7.65 9.5e-11 ***
## Year2009           0.0814     0.1279     0.64  0.526
## Year2010           0.4180     0.2032     2.06  0.044 *
## Year2011          -0.0212     0.1863    -0.11  0.910
## Year2012           0.1265     0.2375     0.53  0.596
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.206, Adjusted R-squared:  -0.00477
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 75 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.503  0.885  0.959  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          1.15e-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"

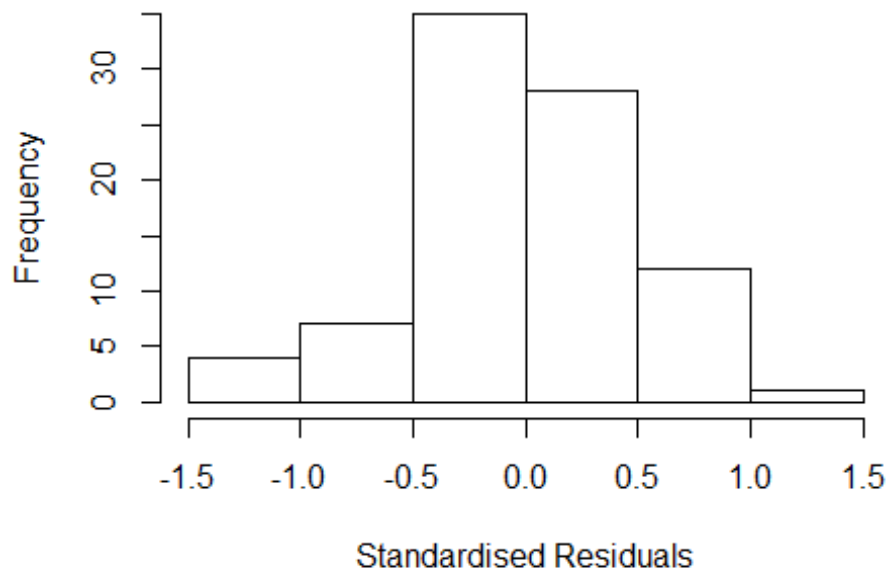
```

```
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

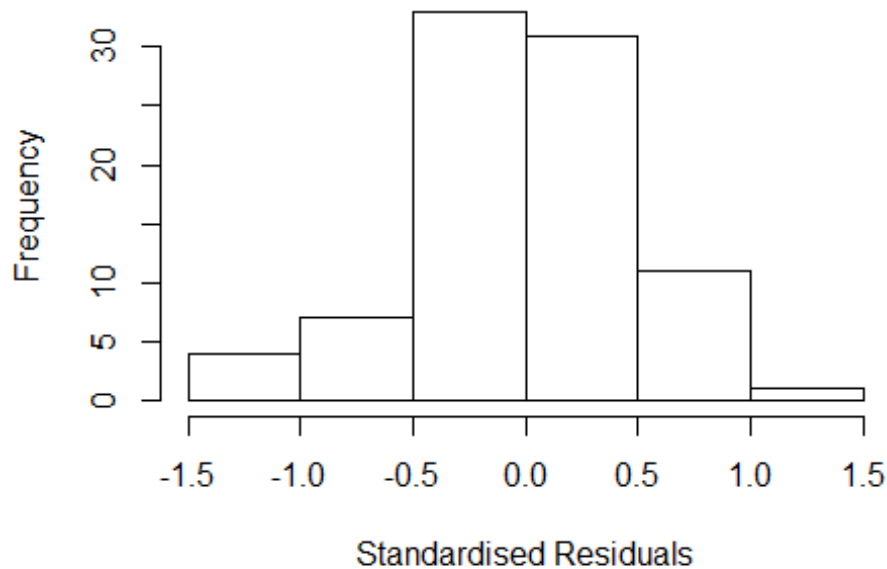
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1          NaN
## Year              NaN 16         NaN
```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.35e+00 -3.05e-01 -2.22e-16 3.14e-01 1.11e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0600 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 0.1731 0.1262 1.37 0.17437
## Year1997 -0.1029 0.2351 -0.44 0.66306
## Year1998 0.3543 0.2145 1.65 0.10320
## Year1999 0.3240 1.1936 0.27 0.78686
## Year2000 -0.0310 0.2700 -0.11 0.90893
## Year2001 0.4085 0.2943 1.39 0.16963
## Year2002 -0.7580 0.0000 -Inf < 2e-16 ***
## Year2003 0.5540 0.1465 3.78 0.00033 ***
## Year2004 0.3213 0.1607 2.00 0.04952 *
## Year2005 0.0750 0.1169 0.64 0.52321
## Year2006 0.0461 0.3020 0.15 0.87907
```

```

## Year2007          0.1695      0.2463      0.69  0.49375
## Year2008          0.6140      0.0811      7.57  1.2e-10 ***
## Year2009          0.1046      0.1243      0.84  0.40312
## Year2010          0.4001      0.2023      1.98  0.05201 .
## Year2011         -0.0402      0.1814     -0.22  0.82518
## Year2012          0.1135      0.2276      0.50  0.61967
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.2,   Adjusted R-squared:  0.00297
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.480  0.880  0.956  0.917  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      1.15e-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 2.872 1      1.695
## Year              2.872 16      1.034

## [1] "List of 0 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
## -1.26e+00 -2.82e-01  1.55e-15  3.22e-01  1.11e+00

```

```

##
## Coefficients:
##           Estimate Std. Error   t value Pr(>|t|)
## (Intercept)  1.06e+00  3.56e-08  2.98e+07 < 2e-16 ***
## LastAuthorFemale1 1.95e-01  1.35e-01  1.44e+00  0.1533
## Year1997      -1.07e-01  2.35e-01 -4.60e-01  0.6502
## Year1998       3.29e-01  1.88e-01  1.75e+00  0.0847 .
## Year1999       3.24e-01  1.27e+00  2.60e-01  0.7989
## Year2000      -2.99e-02  2.72e-01 -1.10e-01  0.9128
## Year2001       4.08e-01  2.96e-01  1.38e+00  0.1719
## Year2002      -7.58e-01  5.62e-08 -1.35e+07 < 2e-16 ***
## Year2003       5.47e-01  1.52e-01  3.60e+00  0.0006 ***
## Year2004       2.56e-01  1.35e-01  1.89e+00  0.0624 .
## Year2005      -5.42e-02  1.22e-01 -4.40e-01  0.6593
## Year2006       1.76e-02  2.83e-01  6.00e-02  0.9505
## Year2007       1.15e-01  2.74e-01  4.20e-01  0.6761
## Year2008       6.44e-01  7.71e-02  8.35e+00  4.7e-12 ***
## Year2009       1.00e-01  1.24e-01  8.10e-01  0.4191
## Year2010       4.69e-01  1.80e-01  2.60e+00  0.0114 *
## Year2011       3.09e-02  1.94e-01  1.60e-01  0.8739
## Year2012       2.02e-01  2.31e-01  8.80e-01  0.3845
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.507
## Multiple R-squared:  0.202, Adjusted R-squared:  0.00578
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.514  0.873   0.956   0.913   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           1.15e-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: 87"
## [1] ""

```

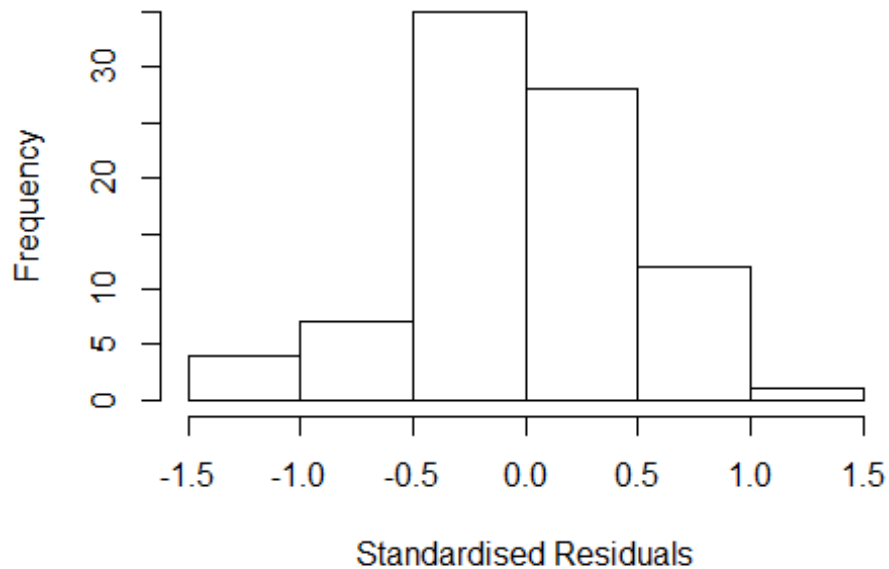
```

## [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
##    2    6    5    4   20    5    1    4    5    9    7    9    4    9    7
## 2011 2012
##    5   11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    2    2   10    1    1    2    4    8    6    8    2    4    5
## 2011 2012
##    3    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    1    2    8    1    1    2    4    8    6    7    2    3    5
## 2011 2012
##    3    7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5"
## [1] "Male first author team size 2018 geometric mean: 5.42883523318981"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

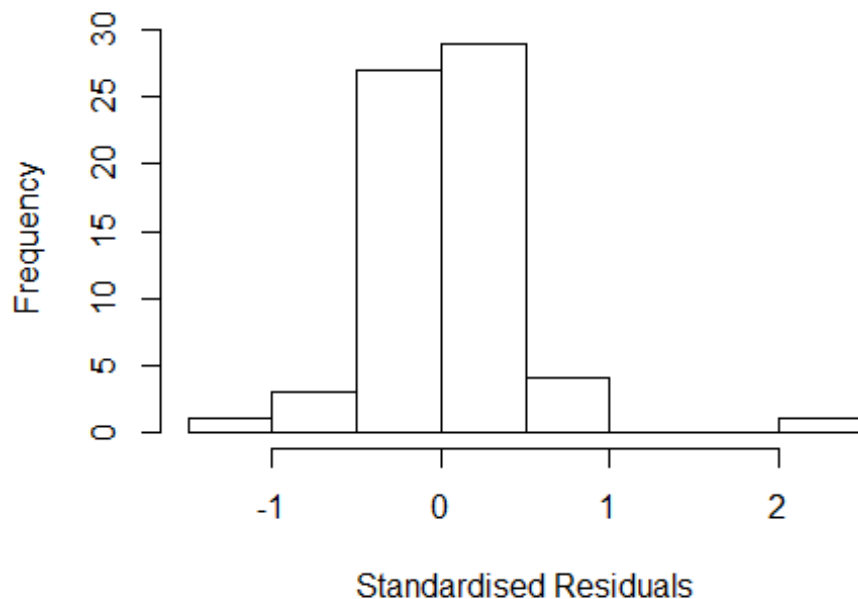

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 5.25305560880753"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF^(1/(2*Df))
## FirstAuthorFemale	-5.891e+00	1	NaN
## LastAuthorFemale	3.419e+00	1	1.849
## UniqueAuthors	1.970e+18	4	193.556
## Year	7.664e+18	16	3.892

Residuals from first and last author and team size



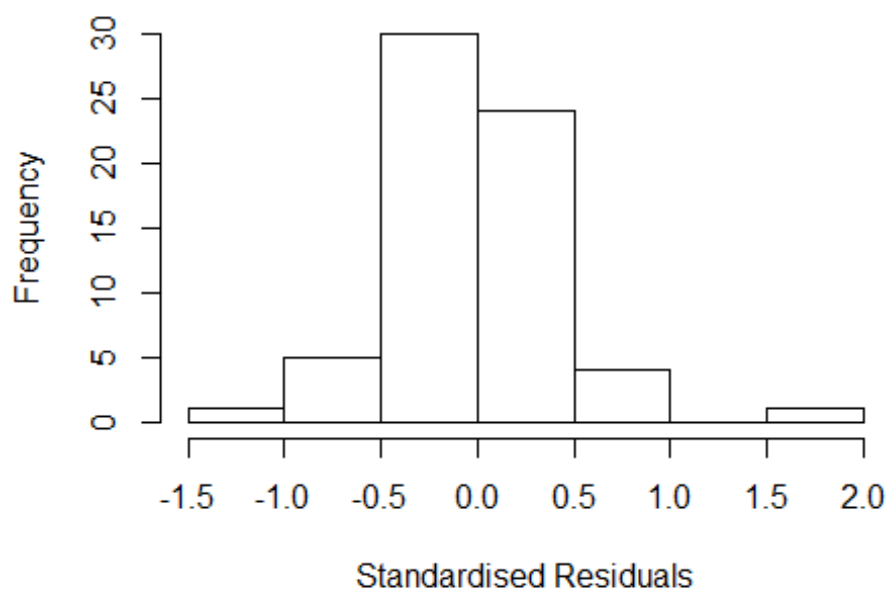
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0613 -0.1216 0.0025 0.1253 2.0417
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9907 0.3497 2.83 0.0071 **
## FirstAuthorFemale1 0.1267 0.1958 0.65 0.5213
## LastAuthorFemale1 -0.1409 0.1101 -1.28 0.2075
## UniqueAuthors2 0.1027 0.2652 0.39 0.7005
## UniqueAuthors3 0.2166 0.3719 0.58 0.5634
## UniqueAuthors4 0.1848 0.2865 0.64 0.5224
## UniqueAuthors5 0.6525 0.3375 1.93 0.0600 .
## Year1997 -0.4086 0.3575 -1.14 0.2595
## Year1998 -0.1517 0.3497 -0.43 0.6666
## Year1999 -0.4918 0.2262 -2.17 0.0354 *
```

```

## Year2000          -1.0762      0.3135    -3.43    0.0014 **
## Year2001          -1.2074      0.2116    -5.70    1.1e-06 ***
## Year2002          -0.6604      0.2116    -3.12    0.0033 **
## Year2003          -0.1498      0.2175    -0.69    0.4946
## Year2004          -0.1618      0.1599    -1.01    0.3174
## Year2005          -0.1459      0.2161    -0.68    0.5032
## Year2006          -0.1787      0.1898    -0.94    0.3518
## Year2007          -0.2982      0.2929    -1.02    0.3145
## Year2008          -0.7223      0.4312    -1.68    0.1014
## Year2009          -0.0673      0.2105    -0.32    0.7509
## Year2010          -0.1280      0.8340    -0.15    0.8788
## Year2011          -0.2276      0.5923    -0.38    0.7027
## Year2012          -0.3991      0.2451    -1.63    0.1109
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.319
## Multiple R-squared:  0.69,   Adjusted R-squared:  0.528
## Convergence in 48 IRWLS iterations
##
## Robustness weights:
## observation 15 is an outlier with |weight| = 0 ( < 0.0015);
## 12 weights are ~= 1. The remaining 52 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.246  0.885  0.975  0.894  0.992  0.999
## Algorithmic parameters:
##           tuning.chi          bb          tuning.psi          refine.tol
##           1.55e+00          5.00e-01          4.69e+00          1.00e-07
##           rel.tol          solve.tol          eps.outlier          eps.x
##           1.00e-07          1.00e-07          1.54e-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 6.728 1 2.594
## LastAuthorFemale -9.353 1 NaN
## Year -5.431 16 NaN

```

Residuals from first and last author



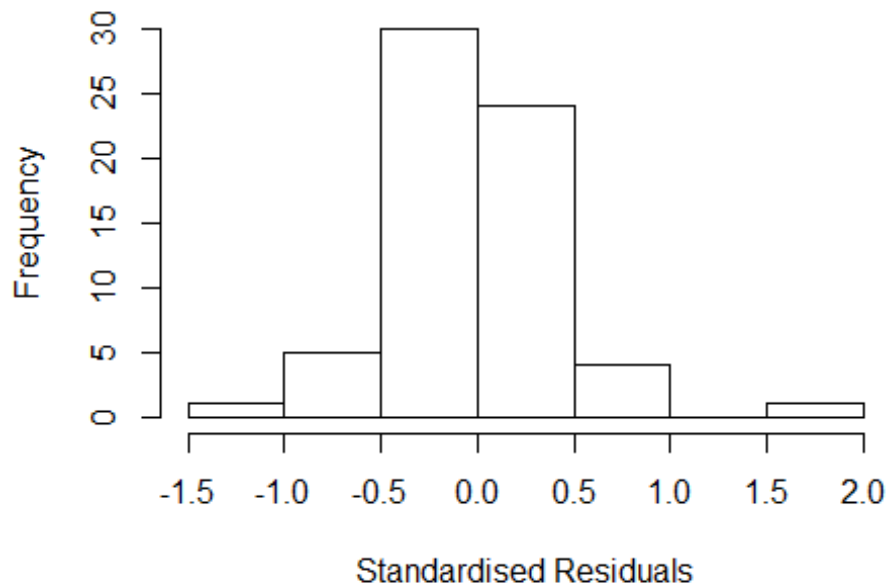
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2648 -0.1424 -0.0181 0.2251 1.8891
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.16131 0.16431 7.07 7.2e-09 ***
## FirstAuthorFemale1 0.00854 0.11269 0.08 0.93995
## LastAuthorFemale1 -0.08061 0.12184 -0.66 0.51153
## Year1997 -0.18346 0.25804 -0.71 0.48068
## Year1998 -0.32231 0.16431 -1.96 0.05588 .
## Year1999 -0.00981 0.16679 -0.06 0.95338
## Year2000 -1.05184 0.19348 -5.44 2.0e-06 ***
## Year2001 -1.16131 0.16431 -7.07 7.2e-09 ***
## Year2002 -0.61431 0.16431 -3.74 0.00051 ***
## Year2003 0.11419 0.17356 0.66 0.51386
## Year2004 -0.11693 0.16756 -0.70 0.48878
## Year2005 0.07158 0.23119 0.31 0.75826
```

```

## Year2006          -0.03518      0.18025      -0.20      0.84613
## Year2007           0.06592      0.29392       0.22      0.82354
## Year2008          -0.24031      0.37668      -0.64      0.52666
## Year2009           0.29671      0.30494       0.97      0.33564
## Year2010           0.10352      0.56624       0.18      0.85574
## Year2011          -0.22620      0.38924      -0.58      0.56398
## Year2012           0.01740      0.23690       0.07      0.94175
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.53,   Adjusted R-squared:  0.347
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## observation 15 is an outlier with |weight| = 0 ( < 0.0015);
## 9 weights are ~1. The remaining 55 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##   0.265  0.910  0.970   0.903   0.992   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      1.54e-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 -6.397 1      NaN
## Year              -6.397 16      NaN

```

Residuals from first author



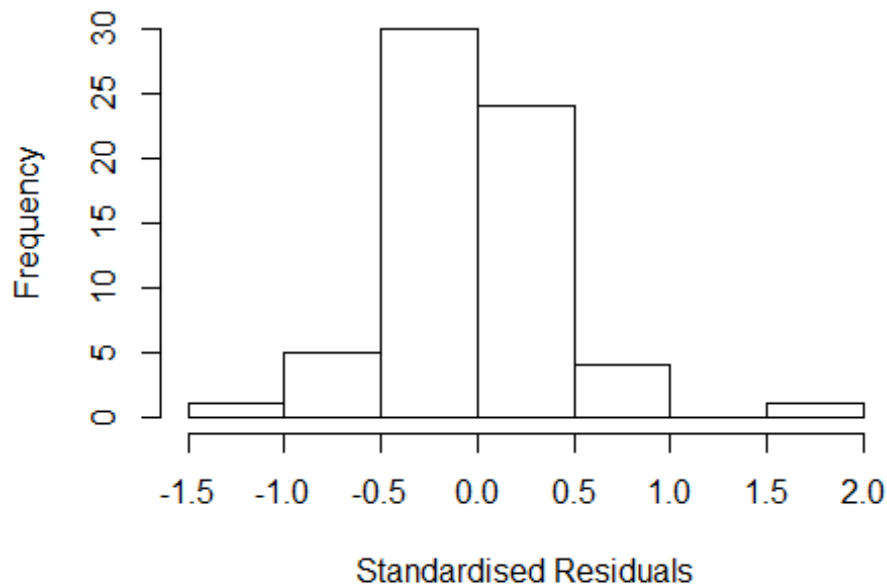
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2903 -0.1471 -0.0203 0.2347 1.8300
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.12100 0.18821 5.96 3.1e-07 ***
## FirstAuthorFemale1 -0.00811 0.11112 -0.07 0.9421
## Year1997 -0.13766 0.27351 -0.50 0.6171
## Year1998 -0.28200 0.18821 -1.50 0.1407
## Year1999 0.03050 0.19039 0.16 0.8734
## Year2000 -1.03302 0.21089 -4.90 1.2e-05 ***
## Year2001 -1.12100 0.18821 -5.96 3.1e-07 ***
## Year2002 -0.57400 0.18821 -3.05 0.0038 **
## Year2003 0.15450 0.19636 0.79 0.4353
## Year2004 -0.09277 0.20051 -0.46 0.6457
## Year2005 0.11085 0.25329 0.44 0.6637
## Year2006 -0.01578 0.20598 -0.08 0.9393
```

```

## Year2007          0.10779    0.31005    0.35    0.7296
## Year2008          -0.20000    0.39457   -0.51    0.6146
## Year2009          0.32792    0.34598    0.95    0.3481
## Year2010          0.16932    0.58933    0.29    0.7751
## Year2011          -0.18427    0.39748   -0.46    0.6451
## Year2012          0.06733    0.25139    0.27    0.7900
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared:  0.539, Adjusted R-squared:  0.372
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## observation 15 is an outlier with |weight| = 0 ( < 0.0015);
## 6 weights are ~= 1. The remaining 58 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.211  0.906  0.966  0.901  0.994  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.54e-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 1.103e+11 1      3.321e+05
## Year              1.103e+11 16      2.214e+00

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2770 -0.1443 -0.0207 0.2189 1.8861
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.16084 0.16436 7.06 6.6e-09 ***
## LastAuthorFemale1 -0.07968 0.11865 -0.67 0.50517
## Year1997 -0.18019 0.25924 -0.70 0.49042
## Year1998 -0.32184 0.16436 -1.96 0.05616 .
## Year1999 -0.00934 0.16684 -0.06 0.95561
## Year2000 -1.04923 0.19223 -5.46 1.8e-06 ***
## Year2001 -1.16084 0.16436 -7.06 6.6e-09 ***
## Year2002 -0.61384 0.16436 -3.73 0.00051 ***
## Year2003 0.11466 0.17362 0.66 0.51220
## Year2004 -0.11454 0.16436 -0.70 0.48930
## Year2005 0.07828 0.20185 0.39 0.69992
## Year2006 -0.03211 0.17297 -0.19 0.85354
```

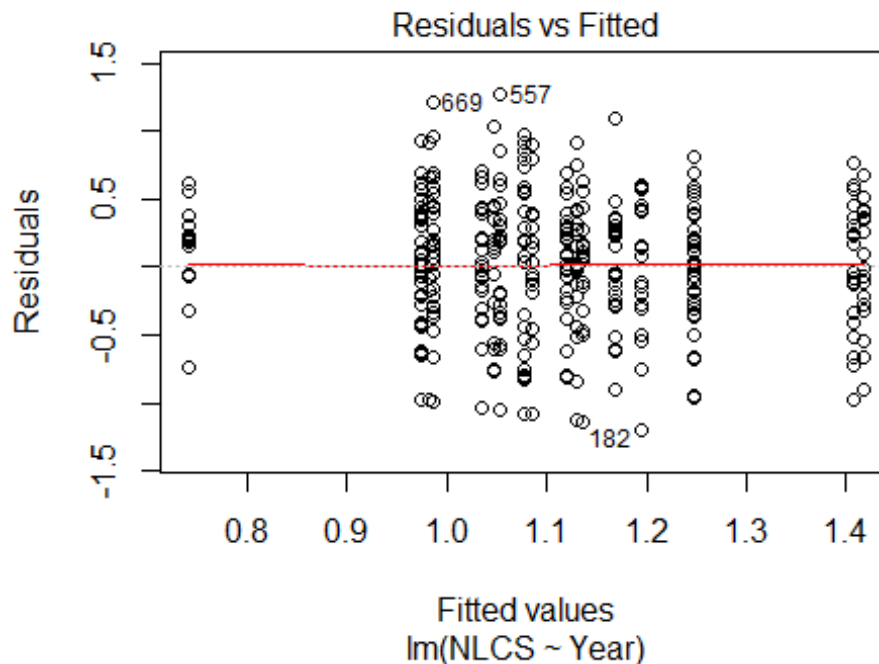


```

## Year2007      0.06723      0.29400      0.23  0.82011
## Year2008     -0.23984      0.38284     -0.63  0.53404
## Year2009      0.30068      0.30153      1.00  0.32378
## Year2010      0.11612      0.57300      0.20  0.84028
## Year2011     -0.22479      0.38904     -0.58  0.56615
## Year2012      0.02268      0.22252      0.10  0.91925
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.381
## Multiple R-squared:  0.536, Adjusted R-squared:  0.369
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## observation 15 is an outlier with |weight| = 0 ( < 0.0015);
## 7 weights are ~ = 1. The remaining 57 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.239  0.907  0.969  0.903  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.54e-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: 65"
## [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
##   14   27   34   44   39   29   27   24   25   34   34   31   53   39   42
## 2011 2012
##   41   40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8   16   18   25   19   18   19   17   16   25   25   20   37   22   30

```

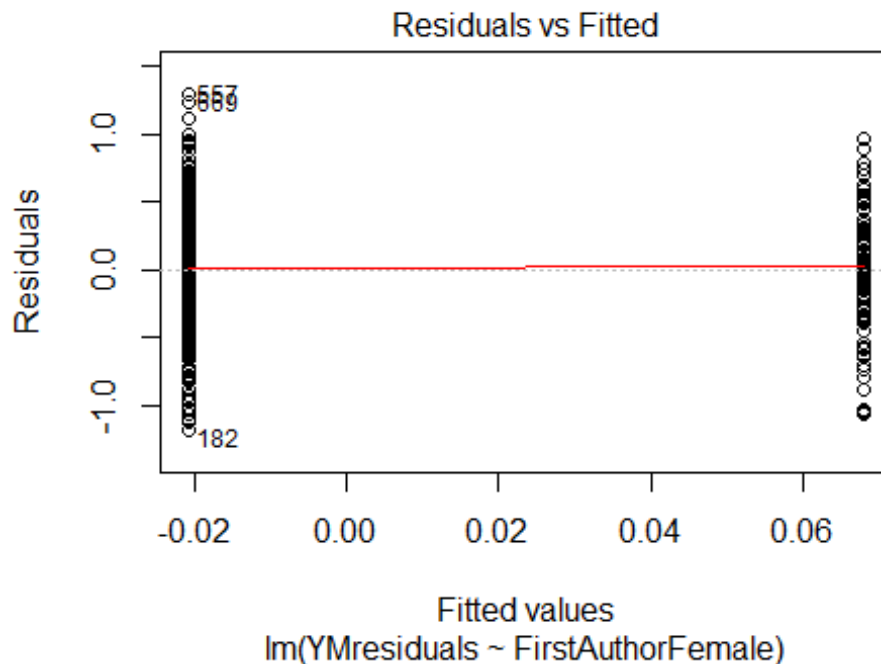
```
## 2011 2012
## 36 36
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 14 16 22 17 17 17 15 13 17 21 19 29 19 27
## 2011 2012
## 30 34
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## 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 = 0.1, df = 1, p-value = 0.7
## [1] "Female first author team size 2018 geometric mean: 5.70932570564503"
## [1] "Male first author team size 2018 geometric mean: 5.06872999546534"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

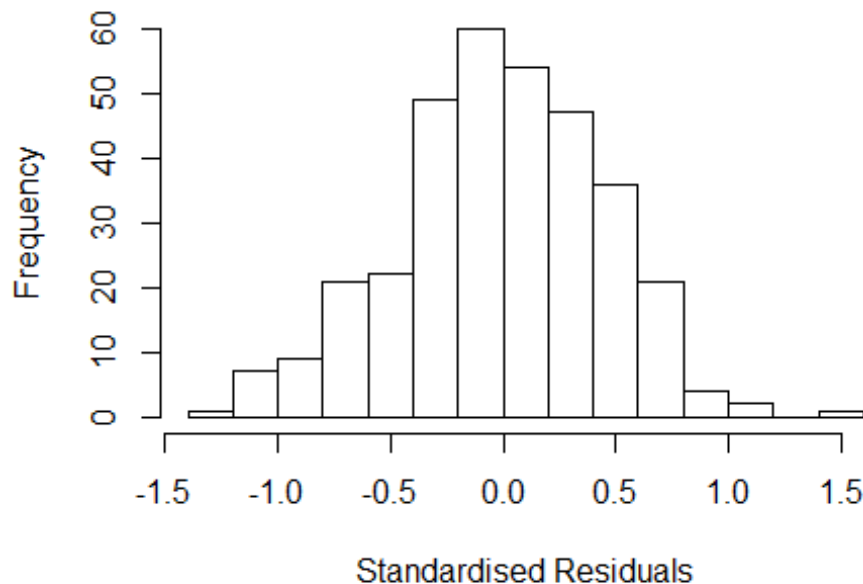
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 60, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.24148278841779"
## [1] "Male last author team size 2018 geometric mean: 5.30106808941112"

## 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 = 28, 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 1.254 1          1.120
## LastAuthorFemale 1.251 1          1.119
## UniqueAuthors    2.221 4          1.105
## Year              2.957 16         1.034
```

Residuals from first and last author and team size



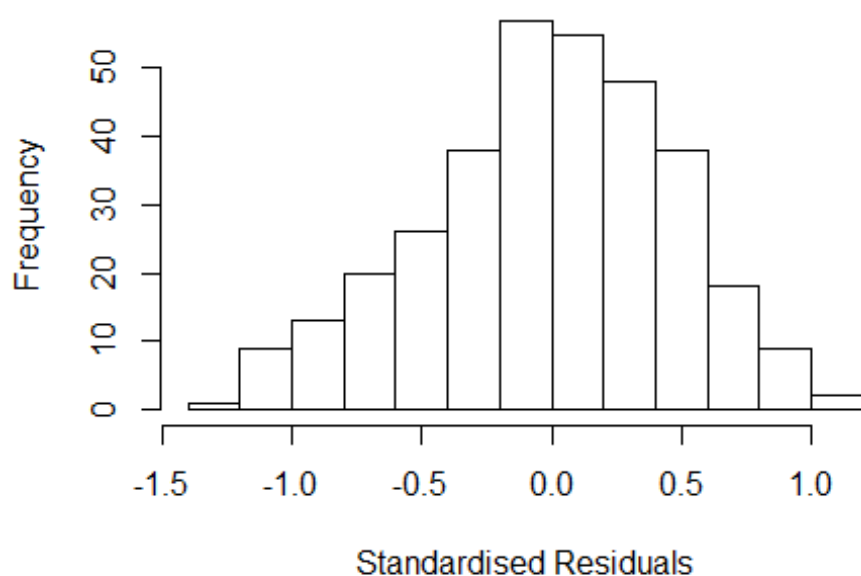
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.24769 -0.27601 -0.00165 0.31659 1.44382
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.82618 0.30579 2.70 0.0073 **
## FirstAuthorFemale1 0.01768 0.06688 0.26 0.7917
## LastAuthorFemale1 0.11878 0.07228 1.64 0.1013
## UniqueAuthors2 -0.07516 0.13449 -0.56 0.5767
## UniqueAuthors3 0.22062 0.13011 1.70 0.0910 .
## UniqueAuthors4 0.18582 0.12319 1.51 0.1325
## UniqueAuthors5 0.33784 0.11299 2.99 0.0030 **
## Year1997 0.12762 0.33302 0.38 0.7018
## Year1998 0.05235 0.33966 0.15 0.8776
## Year1999 0.07042 0.32978 0.21 0.8311
```

```

## Year2000      0.14430      0.33523      0.43      0.6672
## Year2001     -0.24374      0.31591     -0.77      0.4410
## Year2002      0.08366      0.32365      0.26      0.7962
## Year2003      0.31036      0.32901      0.94      0.3463
## Year2004     -0.08441      0.33178     -0.25      0.7993
## Year2005      0.07516      0.32488      0.23      0.8172
## Year2006     -0.13751      0.32156     -0.43      0.6692
## Year2007      0.33876      0.33081      1.02      0.3066
## Year2008      0.14761      0.32135      0.46      0.6463
## Year2009     -0.07480      0.32487     -0.23      0.8181
## Year2010      0.01952      0.32449      0.06      0.9521
## Year2011     -0.00786      0.33896     -0.02      0.9815
## Year2012     -0.05840      0.31685     -0.18      0.8539
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.465
## Multiple R-squared:  0.189, Adjusted R-squared:  0.132
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 304 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.314  0.862  0.956  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      2.99e-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 1.226 1      1.107
## LastAuthorFemale  1.232 1      1.110
## Year              1.508 16      1.013

```

Residuals from first and last author



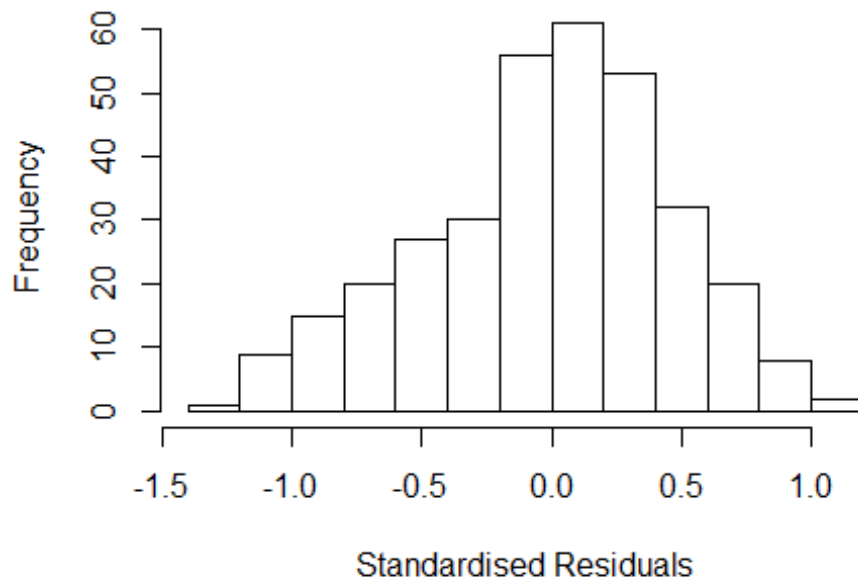
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2191 -0.2980 0.0127 0.3367 1.1693
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.93077 0.35588 2.62 0.0093 **
## FirstAuthorFemale1 0.04596 0.06847 0.67 0.5026
## LastAuthorFemale1 0.10547 0.07287 1.45 0.1488
## Year1997 0.19191 0.37325 0.51 0.6075
## Year1998 0.20910 0.38071 0.55 0.5832
## Year1999 0.14974 0.40495 0.37 0.7118
## Year2000 0.28832 0.38687 0.75 0.4567
## Year2001 -0.22556 0.37068 -0.61 0.5433
## Year2002 0.16926 0.37179 0.46 0.6492
## Year2003 0.47986 0.37340 1.29 0.1997
## Year2004 0.06980 0.38253 0.18 0.8553
## Year2005 0.16991 0.37059 0.46 0.6469
```

```

## Year2006          0.00192    0.37073    0.01    0.9959
## Year2007          0.46310    0.37774    1.23    0.2211
## Year2008          0.30826    0.36418    0.85    0.3979
## Year2009         -0.01456    0.37524   -0.04    0.9691
## Year2010          0.14578    0.36972    0.39    0.6936
## Year2011          0.11227    0.38100    0.29    0.7684
## Year2012          0.01842    0.36770    0.05    0.9601
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.474
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0614
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 45 weights are ~= 1. The remaining 289 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.487  0.861  0.941  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      2.99e-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 1.23 1      1.109
## Year              1.23 16      1.006

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2464 -0.3146 0.0144 0.3382 1.1710
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9317 0.3619 2.57 0.01 *
## FirstAuthorFemale1 0.0554 0.0684 0.81 0.42
## Year1997 0.2074 0.3789 0.55 0.58
## Year1998 0.2179 0.3853 0.57 0.57
## Year1999 0.1504 0.4117 0.37 0.72
## Year2000 0.3146 0.3918 0.80 0.42
## Year2001 -0.2086 0.3764 -0.55 0.58
## Year2002 0.1995 0.3758 0.53 0.60
## Year2003 0.4850 0.3790 1.28 0.20
## Year2004 0.0853 0.3868 0.22 0.83
## Year2005 0.1673 0.3768 0.44 0.66
## Year2006 0.0173 0.3760 0.05 0.96
```

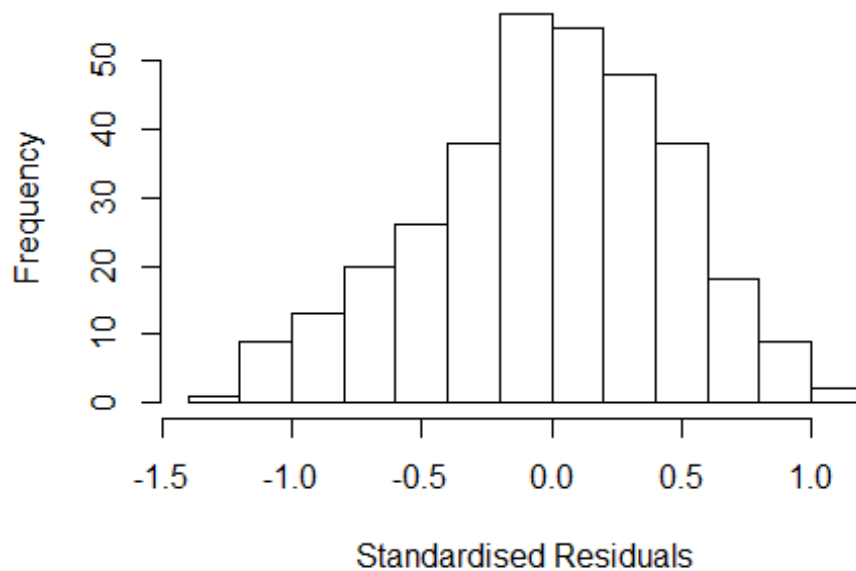


```

## Year2007          0.4632      0.3837      1.21      0.23
## Year2008          0.3240      0.3696      0.88      0.38
## Year2009         -0.0105      0.3808     -0.03      0.98
## Year2010          0.1508      0.3756      0.40      0.69
## Year2011          0.1379      0.3852      0.36      0.72
## Year2012          0.0278      0.3734      0.07      0.94
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.468
## Multiple R-squared:  0.108, Adjusted R-squared:  0.0603
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 37 weights are ~= 1. The remaining 297 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.458  0.855   0.943   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      2.99e-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 1.234 1          1.111
## Year              1.234 16          1.007

```

Residuals from last author



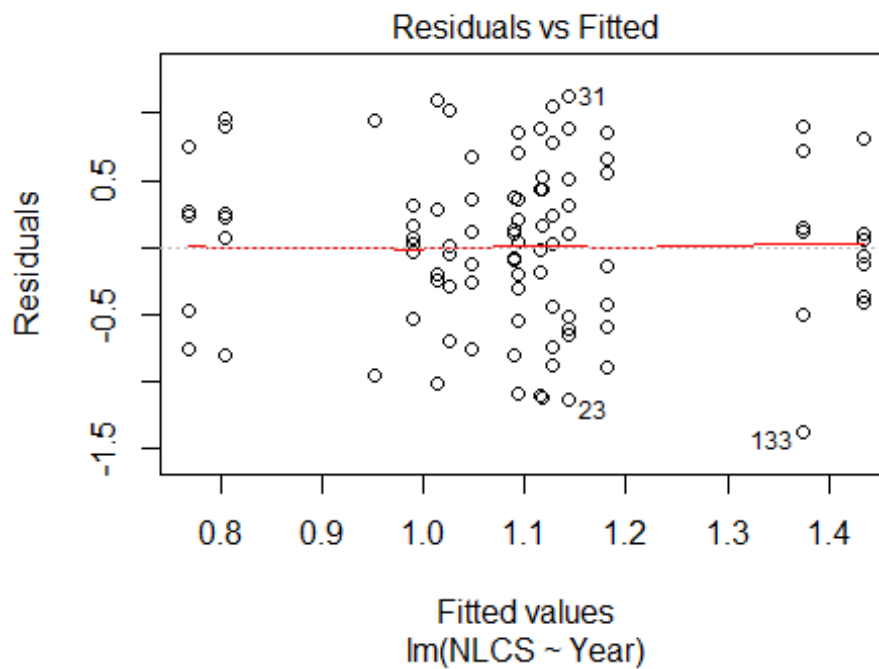
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2275 -0.3091 0.0174 0.3424 1.1638
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9323 0.3522 2.65 0.0085 **
## LastAuthorFemale1 0.1106 0.0730 1.52 0.1307
## Year1997 0.1995 0.3679 0.54 0.5880
## Year1998 0.2077 0.3770 0.55 0.5820
## Year1999 0.1624 0.3991 0.41 0.6843
## Year2000 0.2952 0.3801 0.78 0.4379
## Year2001 -0.2249 0.3663 -0.61 0.5397
## Year2002 0.1756 0.3663 0.48 0.6321
## Year2003 0.4913 0.3672 1.34 0.1819
## Year2004 0.0733 0.3791 0.19 0.8468
## Year2005 0.1739 0.3661 0.48 0.6351
## Year2006 0.0120 0.3647 0.03 0.9738
```

```

## Year2007          0.4763      0.3715      1.28      0.2008
## Year2008          0.3172      0.3587      0.88      0.3772
## Year2009         -0.0115      0.3706     -0.03      0.9752
## Year2010          0.1619      0.3621      0.45      0.6551
## Year2011          0.1329      0.3725      0.36      0.7214
## Year2012          0.0250      0.3631      0.07      0.9451
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.468
## Multiple R-squared:  0.111, Adjusted R-squared:  0.0636
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 47 weights are ~= 1. The remaining 287 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.472  0.854   0.941   0.898   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      2.99e-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: 334"
## [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
##    6   10   16   12   14   13    7    8   15   15    8    9   17   13   17
## 2011 2012
##   16   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    7    6    5    4    7    3    6    7    6    5    5    7    6    9
## 2011 2012

```

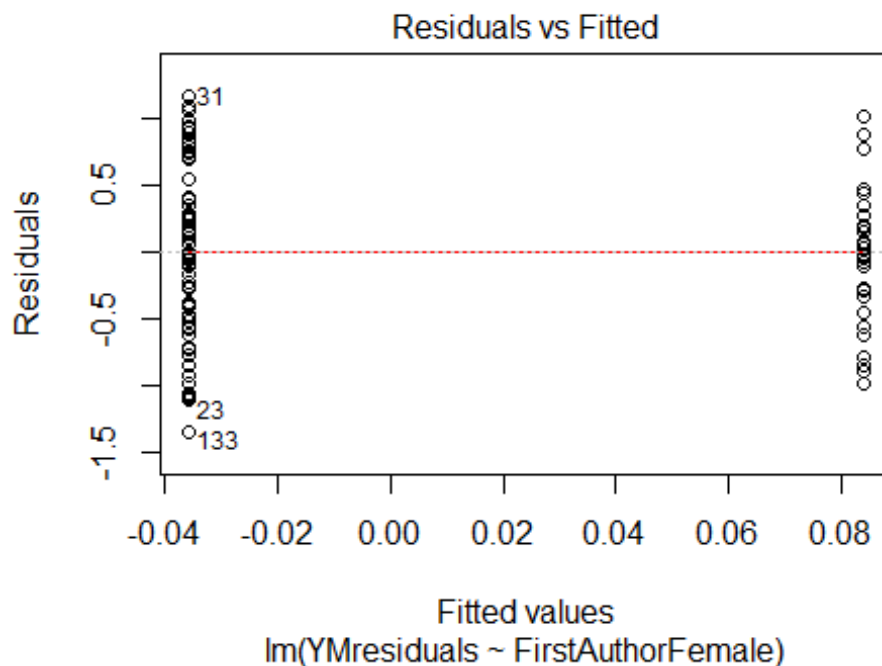
```
##      7      8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      4      6      5      3      5      2      5      4      3      3      4      4      5      8
## 2011 2012
##      7      7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6
```



```
##
## 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: 4.35587717469286"
## [1] "Male first author team size 2018 geometric mean: 2.82842712474619"
## 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 = 6, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.89897948556636"
## [1] "Male last author team size 2018 geometric mean: 3.30975091964687"

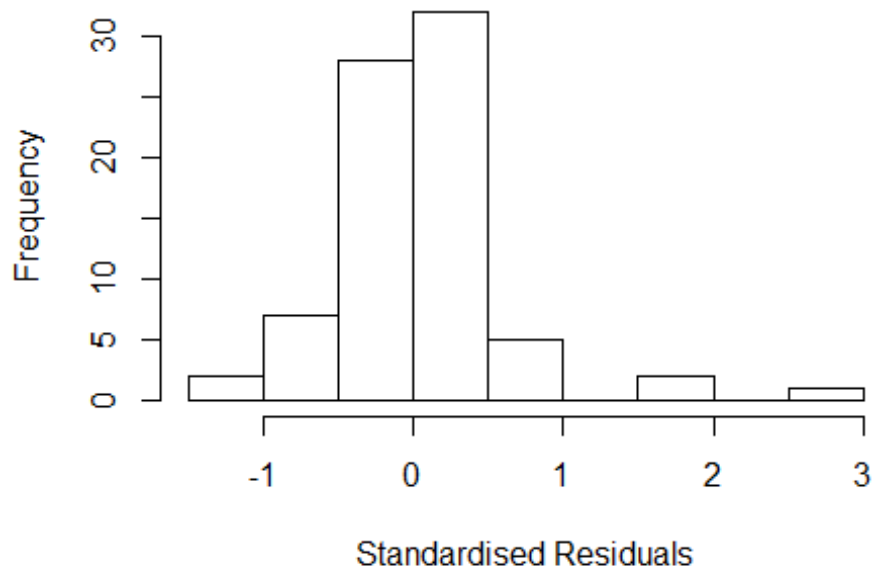
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5.5, 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	4.275	1	2.067
LastAuthorFemale	13.410	1	3.662
UniqueAuthors	616.648	4	2.232
Year	4695.746	16	1.302

Residuals from first and last author and team size



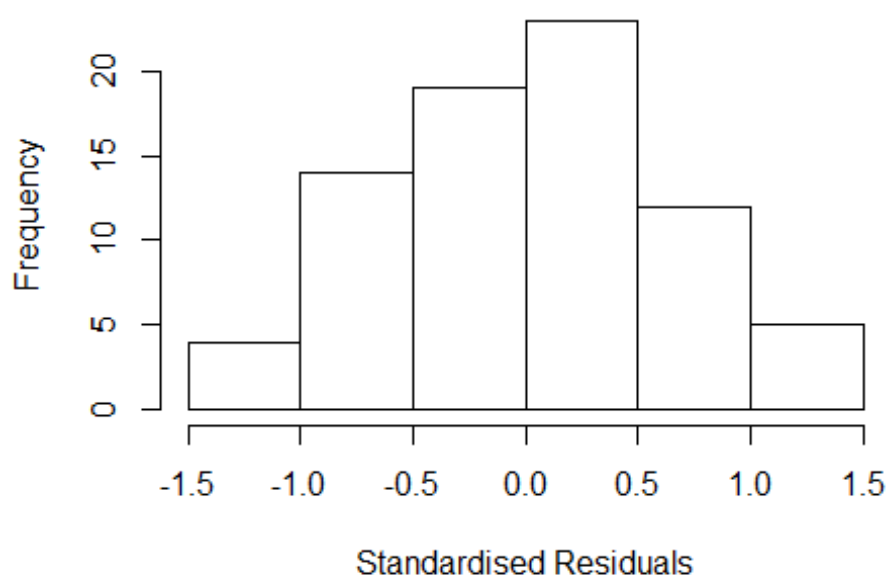
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId NLCS Year OneField Fields residuals
## 144 30344488982 2.27 2005      2705      3      2.672
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0793 -0.2645  0.0174  0.2630  2.6722
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.3797    0.3827    0.99   0.326
## FirstAuthorFemale1 0.3053    0.1393    2.19   0.033 *
## LastAuthorFemale1 0.2329    0.4072    0.57   0.570
## UniqueAuthors2   -0.1438    0.3551   -0.41   0.687
## UniqueAuthors3    0.5548    0.3979    1.39   0.169
## UniqueAuthors4    0.3073    0.6657    0.46   0.646
## UniqueAuthors5    1.1427    0.4640    2.46   0.017 *
## Year1997          0.7954    0.3300    2.41   0.019 *
## Year1998          0.6661    0.8216    0.81   0.421
## Year1999          0.3400    0.3806    0.89   0.376
```

```

## Year2000          -0.1055      0.4203    -0.25      0.803
## Year2001          -0.0305      0.8742    -0.03      0.972
## Year2002           0.2818      0.5726     0.49      0.625
## Year2003           0.0328      0.4453     0.07      0.942
## Year2004           0.4320      0.6618     0.65      0.517
## Year2005          -0.6381      0.4278    -1.49      0.142
## Year2006          -0.2087      0.3286    -0.64      0.528
## Year2007           0.3381      0.3581     0.94      0.349
## Year2008          -0.2752      0.4810    -0.57      0.570
## Year2009           0.4733      0.3588     1.32      0.193
## Year2010           0.0421      0.4469     0.09      0.925
## Year2011           0.1448      0.5021     0.29      0.774
## Year2012          -0.0313      0.3782    -0.08      0.934
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared:  0.476, Adjusted R-squared:  0.263
## Convergence in 35 IRWLS iterations
##
## Robustness weights:
## observation 39 is an outlier with |weight| = 0 ( < 0.0013);
## 9 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.271  0.906  0.965  0.903  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          1.30e-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 3.745 1          1.935
## LastAuthorFemale  2.746 1          1.657
## Year              10.283 16          1.076

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.2011 -0.4864 0.0343 0.3262 1.1565
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.95100 1.14617 0.83 0.41
## FirstAuthorFemale1 0.24678 0.22491 1.10 0.28
## LastAuthorFemale1 0.06354 0.31741 0.20 0.84
## Year1997 0.26456 1.14932 0.23 0.82
## Year1998 0.18076 1.23377 0.15 0.88
## Year1999 -0.06245 1.19256 -0.05 0.96
## Year2000 -0.04243 1.22728 -0.03 0.97
## Year2001 0.02185 1.21011 0.02 0.99
## Year2002 0.14150 1.23506 0.11 0.91
## Year2003 -0.11903 1.16981 -0.10 0.92
## Year2004 0.32474 1.15198 0.28 0.78
## Year2005 0.25010 1.40830 0.18 0.86
```

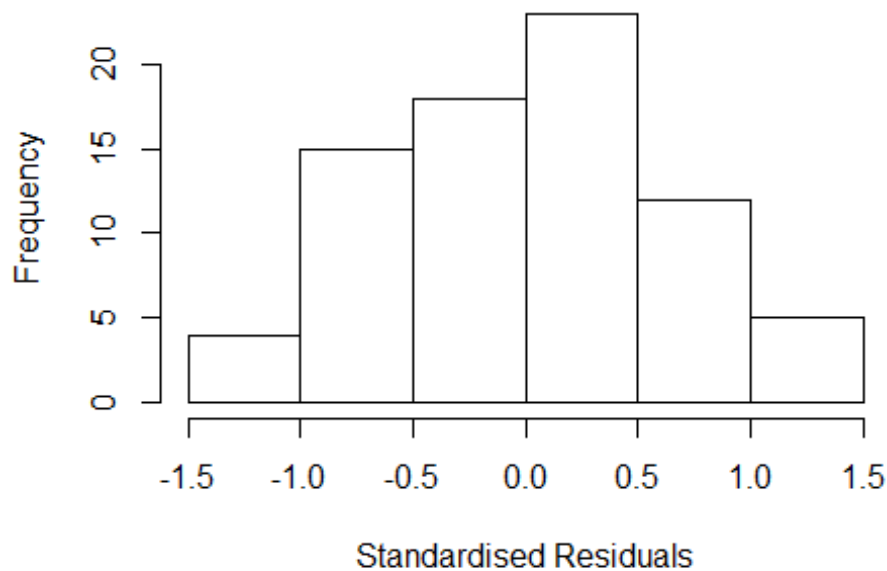


```

## Year2006      -0.19448    1.22773   -0.16    0.87
## Year2007      0.22163    1.21070    0.18    0.86
## Year2008     -0.00566    1.16378    0.00    1.00
## Year2009     -0.07035    1.16310   -0.06    0.95
## Year2010      0.08593    1.16583    0.07    0.94
## Year2011     -0.16032    1.20759   -0.13    0.89
## Year2012     -0.06606    1.17642   -0.06    0.96
##
## Robust residual standard error: 0.663
## Multiple R-squared:  0.0681, Adjusted R-squared:  -0.221
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.724  0.866  0.946  0.919  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      1.30e-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 3.228 1          1.797
## Year              3.228 16          1.037

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.1946 -0.4974 0.0348 0.3524 1.1525
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9510 1.0578 0.90 0.37
## FirstAuthorFemale1 0.2488 0.2161 1.15 0.25
## Year1997 0.2646 1.0612 0.25 0.80
## Year1998 0.1826 1.1425 0.16 0.87
## Year1999 -0.0585 1.0987 -0.05 0.96
## Year2000 -0.0232 1.1432 -0.02 0.98
## Year2001 0.0536 1.1020 0.05 0.96
## Year2002 0.1415 1.1504 0.12 0.90
## Year2003 -0.0970 1.0777 -0.09 0.93
## Year2004 0.3242 1.0635 0.30 0.76
## Year2005 0.2436 1.3162 0.19 0.85
## Year2006 -0.1953 1.1413 -0.17 0.86
```

```

## Year2007          0.2395      1.1132      0.22      0.83
## Year2008          0.0098      1.0720      0.01      0.99
## Year2009         -0.0726      1.0791     -0.07      0.95
## Year2010          0.0874      1.0810      0.08      0.94
## Year2011         -0.1465      1.1204     -0.13      0.90
## Year2012         -0.0712      1.0970     -0.06      0.95
##
## Robust residual standard error: 0.703
## Multiple R-squared:  0.0663, Adjusted R-squared:  -0.203
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.754  0.880  0.951  0.925  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      1.30e-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 2.95  1          1.718
## Year              2.95 16          1.034

## [1] "List of 0 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
## -1.3428 -0.4996  0.0409  0.3304  1.1347
##
## Coefficients:

```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9510      1.0704   0.89   0.38
## LastAuthorFemale1 0.0892      0.3288   0.27   0.79
## Year1997          0.2646      1.0737   0.25   0.81
## Year1998          0.1823      1.1552   0.16   0.88
## Year1999          0.0451      1.0994   0.04   0.97
## Year2000          0.0294      1.1953   0.02   0.98
## Year2001          0.0104      1.1346   0.01   0.99
## Year2002          0.1415      1.1624   0.12   0.90
## Year2003          0.0166      1.0860   0.02   0.99
## Year2004          0.3875      1.0754   0.36   0.72
## Year2005          0.3918      1.3114   0.30   0.77
## Year2006         -0.0909      1.1486  -0.08   0.94
## Year2007          0.4137      1.0986   0.38   0.71
## Year2008          0.0413      1.0988   0.04   0.97
## Year2009         -0.0332      1.0844  -0.03   0.98
## Year2010          0.1459      1.0977   0.13   0.89
## Year2011          0.0272      1.1230   0.02   0.98
## Year2012         -0.0170      1.1080  -0.02   0.99
##
## Robust residual standard error: 0.696
## Multiple R-squared:  0.049, Adjusted R-squared:  -0.225
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.690  0.861  0.941  0.922  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      1.30e-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: 77"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2707"
## [1] "#####"

```

```

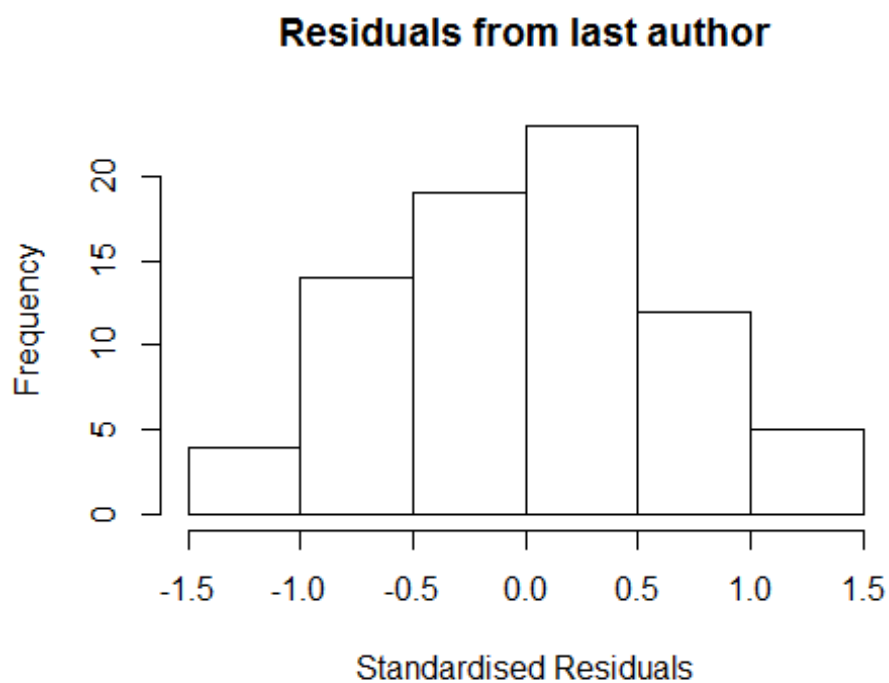
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    1    5    4    3    4    1    7    9    5    6    3    6    2    11
## 2012
##    6
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    3    0    4    4    2    2    1    2    8    4    5    3    5    1    11
## 2012
##    5
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    0    4    4    1    2    1    1    8    4    5    3    3    1    10
## 2012
##    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.77976314968462"
## [1] "Male first author team size 2018 geometric mean: 3.31445401733999"

## 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 = 9, p-value = 0.8
## 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: 4.18856854179052"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

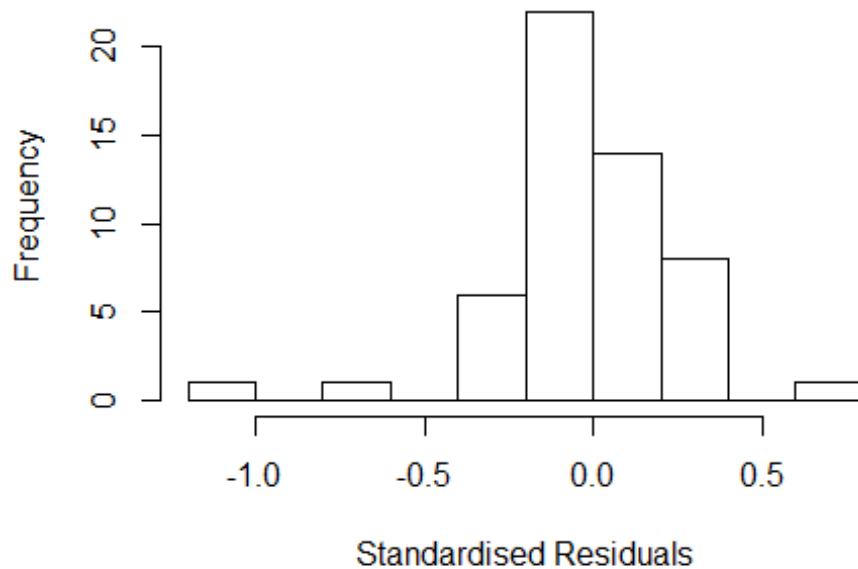
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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	7.996e+15	1	8.942e+07
## LastAuthorFemale	9.163e+01	1	9.572e+00
## UniqueAuthors	5.489e+31	4	9.278e+03
## Year	4.947e+32	14	1.471e+01

Residuals from first and last author and team size



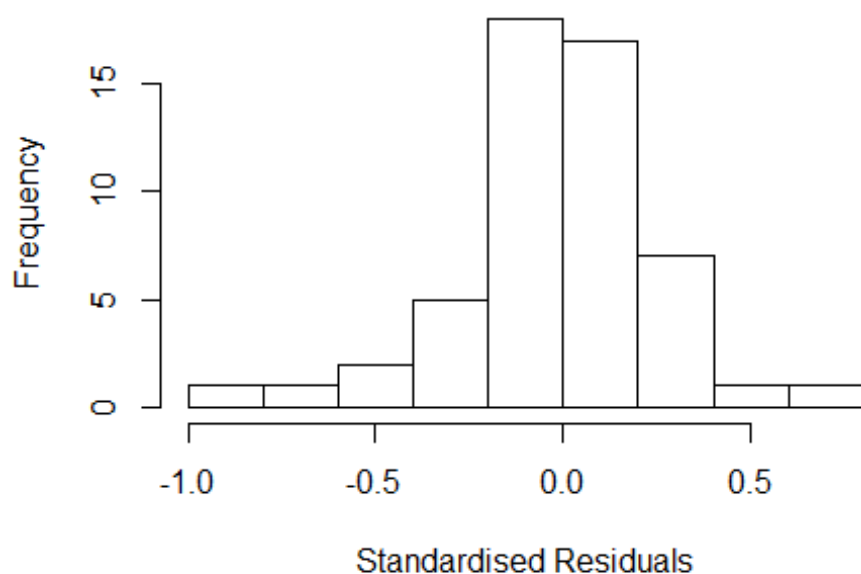
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06e+00 -1.25e-01 3.33e-16 9.06e-02 6.27e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02225 0.13822 7.40 2.1e-08 ***
## FirstAuthorFemale1 -0.02067 0.10245 -0.20 0.8414
## LastAuthorFemale1 0.04160 0.09122 0.46 0.6514
## UniqueAuthors2 0.19451 0.21325 0.91 0.3685
## UniqueAuthors3 0.21545 0.13477 1.60 0.1197
## UniqueAuthors4 0.30007 0.11320 2.65 0.0124 *
## UniqueAuthors5 0.29146 0.12974 2.25 0.0317 *
## Year1999 -0.04291 0.16880 -0.25 0.8009
## Year2000 -0.06580 0.13238 -0.50 0.6226
## Year2001 -0.11332 0.08479 -1.34 0.1908
```

```

## Year2002          -0.05717      0.10715      -0.53      0.5973
## Year2003           0.93097      0.13624       6.83      1.0e-07 ***
## Year2004          -0.21004      0.13659      -1.54      0.1339
## Year2005           0.00423      0.09941       0.04      0.9664
## Year2006          -0.17783      0.12340      -1.44      0.1593
## Year2007           0.05689      0.14634       0.39      0.7001
## Year2008          -0.17036      0.20497      -0.83      0.4121
## Year2009          -0.18582      0.16961      -1.10      0.2814
## Year2010          -0.06004      0.13659      -0.44      0.6632
## Year2011          -0.17324      0.21271      -0.81      0.4214
## Year2012          -0.39683      0.13052      -3.04      0.0047 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.247
## Multiple R-squared:  0.52,   Adjusted R-squared:  0.22
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
##  10 weights are ~= 1. The remaining 43 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0235 0.8780 0.9650 0.8970 0.9900 0.9980
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.89e-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 4.558e+15 1      6.751e+07
## LastAuthorFemale  1.291e+01 1      3.593e+00
## Year              -2.117e+16 14      NaN

```


Residuals from first and last author



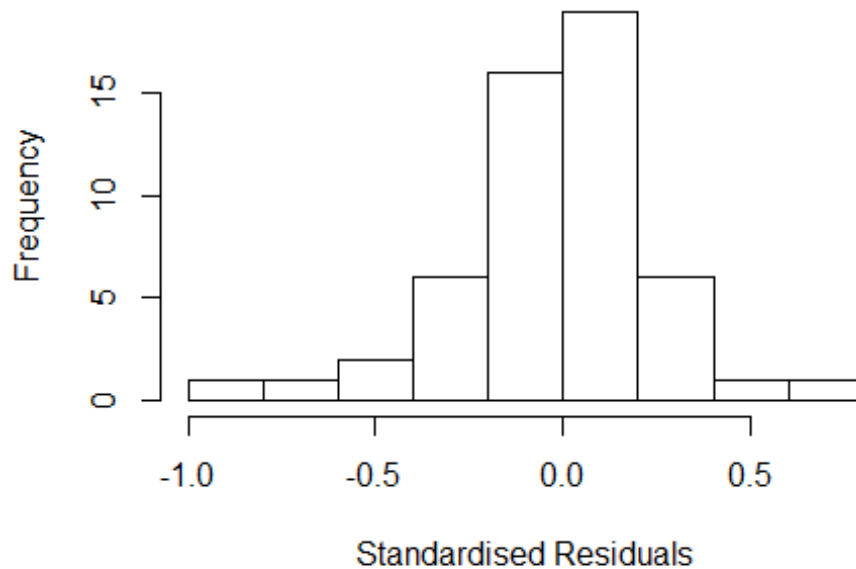
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.69e-01 -1.00e-01 7.49e-16 1.37e-01 6.80e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2908 0.0782 16.51 < 2e-16 ***
## FirstAuthorFemale1 -0.0101 0.0771 -0.13 0.897
## LastAuthorFemale1 -0.0115 0.1158 -0.10 0.921
## Year1999 -0.1308 0.2035 -0.64 0.524
## Year2000 -0.0751 0.1268 -0.59 0.557
## Year2001 -0.0818 0.0782 -1.05 0.302
## Year2002 -0.0733 0.0981 -0.75 0.460
## Year2003 0.8673 0.0772 11.23 2.5e-13 ***
## Year2004 -0.1977 0.0772 -2.56 0.015 *
## Year2005 -0.0372 0.0863 -0.43 0.669
## Year2006 -0.2177 0.1569 -1.39 0.174
## Year2007 0.0341 0.0698 0.49 0.629
```

```

## Year2008          -0.1803      0.2164   -0.83    0.410
## Year2009          -0.2529      0.2686   -0.94    0.353
## Year2010          -0.0477      0.0772   -0.62    0.540
## Year2011          -0.3000      0.1501   -2.00    0.053 .
## Year2012          -0.4708      0.0759   -6.20    3.7e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.281
## Multiple R-squared:  0.405, Adjusted R-squared:  0.14
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.211  0.899  0.973  0.909  0.991  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      1.89e-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 1.774e+14 1      1.332e+07
## Year              1.774e+14 14      3.228e+00

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.98067 -0.10456  0.00239  0.13733  0.67892
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2857     0.0502   25.63 < 2e-16 ***
## FirstAuthorFemale1 -0.0114     0.0804   -0.14  0.8878
## Year1999          -0.1282     0.1891   -0.68  0.5020
## Year2000          -0.0719     0.1167   -0.62  0.5415
## Year2001          -0.0767     0.0502   -1.53  0.1348
## Year2002          -0.0675     0.0769   -0.88  0.3856
## Year2003           0.8737     0.0502   17.42 < 2e-16 ***
## Year2004          -0.1913     0.0502   -3.81  0.0005 ***
## Year2005          -0.0358     0.0827   -0.43  0.6673
## Year2006          -0.2120     0.1419   -1.49  0.1437
## Year2007           0.0333     0.0696    0.48  0.6356
## Year2008          -0.1735     0.2057   -0.84  0.4044
```

```

## Year2009          -0.2559      0.2468   -1.04    0.3066
## Year2010          -0.0413      0.0502   -0.82    0.4157
## Year2011          -0.2936      0.1354   -2.17    0.0366 *
## Year2012          -0.4681      0.0640   -7.32    1.1e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.277
## Multiple R-squared:  0.407, Adjusted R-squared:  0.167
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 45 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.186  0.898  0.974  0.909  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      1.89e-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 5.315 1          2.305
## Year            5.315 14          1.061

## [1] "List of 0 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
## -0.97540 -0.10425  0.00234  0.13094  0.68394
##
## Coefficients:

```

```

##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2858    0.0688   18.68 < 2e-16 ***
## LastAuthorFemale1 -0.0117    0.1144   -0.10  0.9193
## Year1999         -0.1308    0.2077   -0.63  0.5326
## Year2000         -0.0730    0.1289   -0.57  0.5746
## Year2001         -0.0768    0.0688   -1.12  0.2716
## Year2002         -0.0733    0.1019   -0.72  0.4763
## Year2003          0.8622    0.0688   12.52  7.1e-15 ***
## Year2004         -0.2028    0.0688   -2.95  0.0055 **
## Year2005         -0.0359    0.0873   -0.41  0.6832
## Year2006         -0.2157    0.1579   -1.37  0.1802
## Year2007          0.0331    0.0703    0.47  0.6410
## Year2008         -0.1849    0.2147   -0.86  0.3947
## Year2009         -0.2495    0.2700   -0.92  0.3614
## Year2010         -0.0528    0.0688   -0.77  0.4477
## Year2011         -0.2988    0.1361   -2.20  0.0345 *
## Year2012         -0.4682    0.0721   -6.50  1.3e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.276
## Multiple R-squared:  0.408, Adjusted R-squared:  0.168
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 45 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.185  0.897  0.974  0.908  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      1.89e-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: 53"
## [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
##    11    5    7    4   11   11    4    5    7    5    2   10    8   12   21
## 2011 2012
##    12   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     3    0    1    2    8    3    3    5    5    4    1   10    5    7   13
## 2011 2012
##     6    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     3    0    1    2    6    3    3    5    5    4    1    8    5    5   13
## 2011 2012
##     5    7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male first author team size 2018 geometric mean: 1.8881750225898"

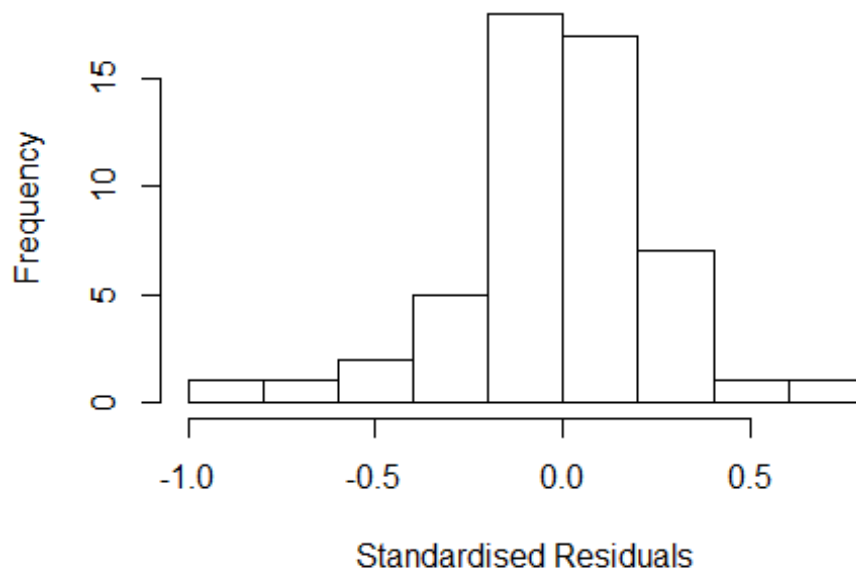
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.21336383940064"
## [1] "Male last author team size 2018 geometric mean: 2.04767251107922"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

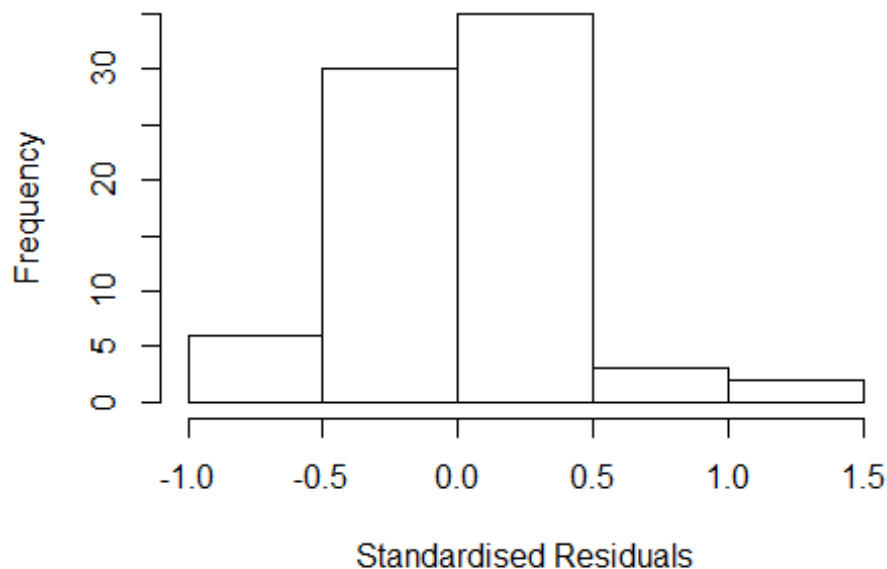
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, 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 1.530e+13  1      3.911e+06
## LastAuthorFemale  5.448e+00  1      2.334e+00
## UniqueAuthors    1.927e+16  4      1.085e+02
## Year             1.045e+17 15      3.692e+00
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -0.9461 -0.2326 0.0291 0.2037 1.3726
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1806 0.1822 6.48 2.9e-08 ***
## FirstAuthorFemale1 0.1087 0.1054 1.03 0.3070
## LastAuthorFemale1 0.0708 0.1775 0.40 0.6914
## UniqueAuthors2 0.2375 0.1636 1.45 0.1523
## UniqueAuthors3 0.4963 0.2471 2.01 0.0497 *
## UniqueAuthors4 0.2308 0.2051 1.13 0.2654
## UniqueAuthors5 0.1442 0.2323 0.62 0.5374
## Year1998 -0.5428 0.3161 -1.72 0.0917 .
## Year1999 -0.2868 0.2707 -1.06 0.2942
## Year2000 -0.1829 0.1863 -0.98 0.3307
```

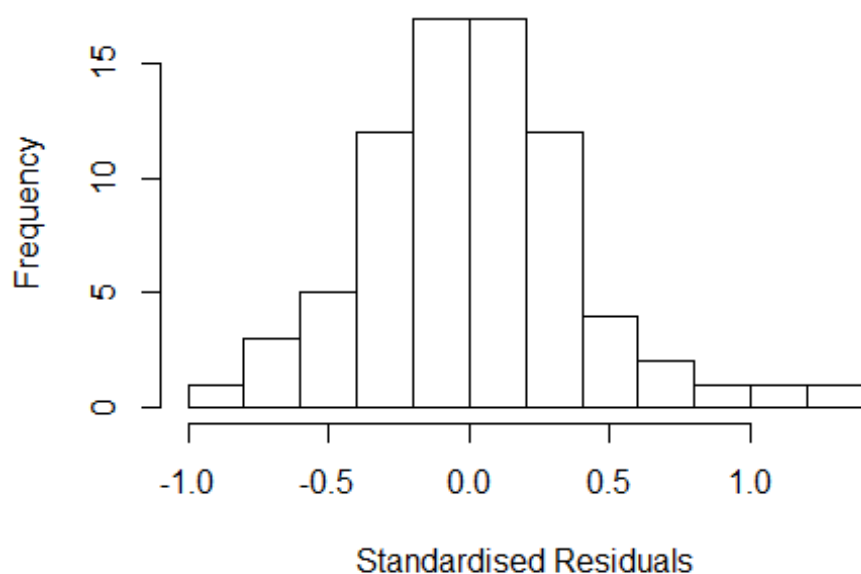


```

## Year2001          -0.4123      0.2806   -1.47    0.1475
## Year2002          -0.4930      0.1979   -2.49    0.0159 *
## Year2003          -0.3547      0.2205   -1.61    0.1136
## Year2004          -0.2809      0.2377   -1.18    0.2424
## Year2005          -0.4475      0.2641   -1.69    0.0959 .
## Year2006          -0.6747      0.2486   -2.71    0.0089 **
## Year2007          -0.2451      0.2735   -0.90    0.3742
## Year2008          -0.4720      0.3301   -1.43    0.1586
## Year2009          -0.9248      0.2677   -3.45    0.0011 **
## Year2010          -0.3261      0.3094   -1.05    0.2965
## Year2011          -0.1511      0.2447   -0.62    0.5397
## Year2012          -0.5280      0.2841   -1.86    0.0685 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.39
## Multiple R-squared:  0.283, Adjusted R-squared:  0.00381
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.189  0.911  0.968  0.916  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.32e-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 -3.108e+15  1      NaN
## LastAuthorFemale  -4.275e+00  1      NaN
## Year              -1.218e+16 15      NaN

```

Residuals from first and last author



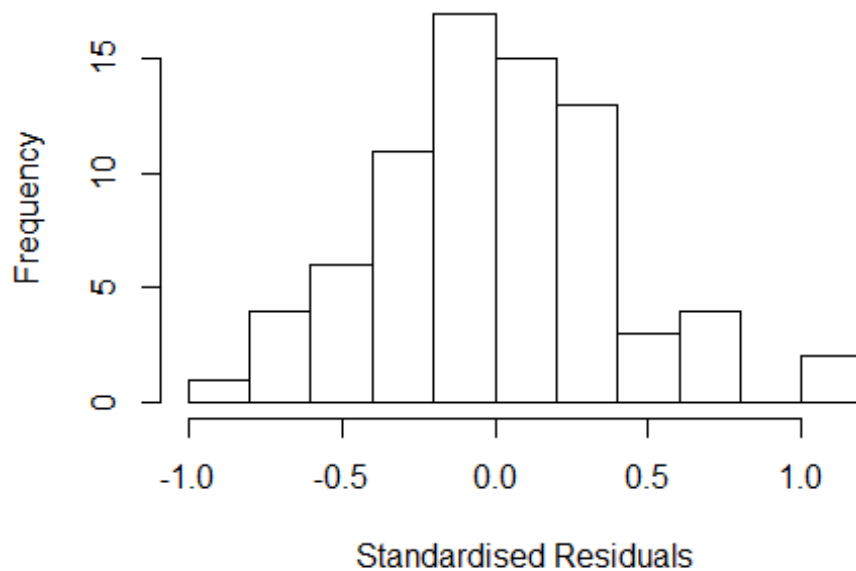
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.91149 -0.24104  0.00158  0.21381  1.21714
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.13941    0.16945   6.72 8.5e-09 ***
## FirstAuthorFemale1  0.13347    0.09737   1.37  0.1757
## LastAuthorFemale1  0.10539    0.15406   0.68  0.4966
## Year1998          -0.00541    0.16945  -0.03  0.9746
## Year1999          -0.10141    0.17973  -0.56  0.5748
## Year2000          -0.04637    0.18859  -0.25  0.8066
## Year2001          -0.06322    0.27485  -0.23  0.8189
## Year2002          -0.39486    0.24469  -1.61  0.1120
## Year2003          -0.14744    0.19770  -0.75  0.4588
## Year2004          -0.04668    0.22855  -0.20  0.8389
## Year2005          -0.16148    0.16394  -0.98  0.3287
## Year2006          -0.42088    0.16798  -2.51  0.0151 *
```

```

## Year2007          -0.02567      0.20572    -0.12    0.9011
## Year2008          -0.22792      0.28881    -0.79    0.4332
## Year2009          -0.67300      0.20705    -3.25    0.0019 **
## Year2010          -0.05896      0.23648    -0.25    0.8040
## Year2011           0.09348      0.24123     0.39    0.6998
## Year2012          -0.29356      0.20724    -1.42    0.1620
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.205, Adjusted R-squared:  -0.0282
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.320  0.890  0.962  0.908  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      1.32e-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 5.336e+14 1      2.310e+07
## Year              5.336e+14 15      3.097e+00

```

Residuals from first author



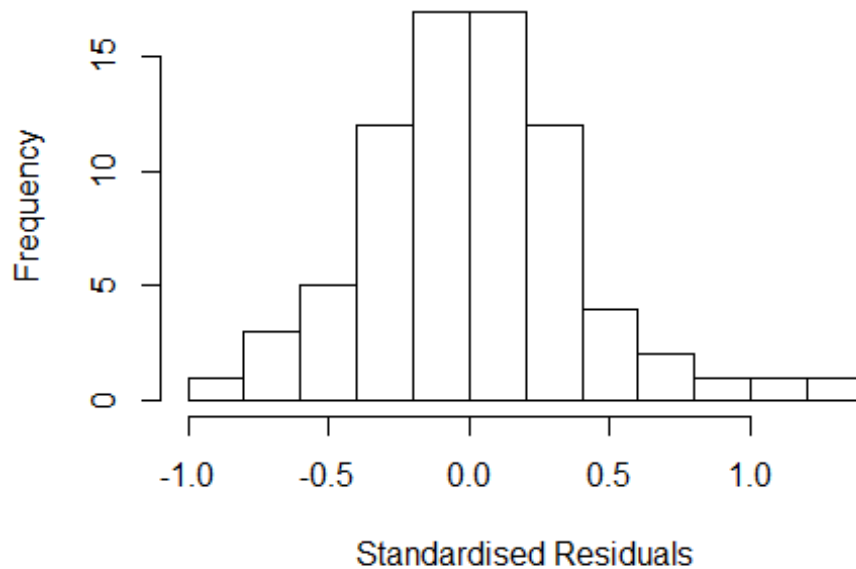
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.77e-01 -2.69e-01 -3.78e-16 2.19e-01 1.20e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1956 0.1641 7.28 8.9e-10 ***
## FirstAuthorFemale1 0.1573 0.1046 1.50 0.13779
## Year1998 -0.0616 0.1641 -0.38 0.70863
## Year1999 -0.1576 0.1747 -0.90 0.37057
## Year2000 -0.0769 0.2060 -0.37 0.71030
## Year2001 -0.1207 0.2685 -0.45 0.65466
## Year2002 -0.4264 0.2265 -1.88 0.06473 .
## Year2003 -0.2143 0.1798 -1.19 0.23802
## Year2004 -0.1139 0.2136 -0.53 0.59594
## Year2005 -0.1729 0.1807 -0.96 0.34269
## Year2006 -0.5009 0.1405 -3.56 0.00073 ***
## Year2007 -0.0842 0.1993 -0.42 0.67417
```

```

## Year2008          -0.2190      0.3096   -0.71  0.48211
## Year2009          -0.7273      0.1948   -3.73  0.00043 ***
## Year2010          -0.0827      0.2408   -0.34  0.73243
## Year2011           0.0393      0.2013    0.20  0.84591
## Year2012          -0.3320      0.2251   -1.48  0.14551
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.408
## Multiple R-squared:  0.192, Adjusted R-squared:  -0.0277
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.367  0.903   0.961   0.916   0.992   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.32e-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 3.509 1          1.873
## Year            3.509 15          1.043

```

Residuals from last author



```
## [1] "List of 0 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
## -0.91788 -0.23577 0.00538 0.22438 1.16325
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1969 0.1876 6.38 3e-08 ***
## LastAuthorFemale1 0.1562 0.1567 1.00 0.3229
## Year1998 -0.0629 0.1876 -0.34 0.7387
## Year1999 -0.1589 0.1969 -0.81 0.4231
## Year2000 -0.0958 0.2074 -0.46 0.6459
## Year2001 -0.1201 0.2845 -0.42 0.6746
## Year2002 -0.4273 0.2350 -1.82 0.0740 .
## Year2003 -0.1456 0.2183 -0.67 0.5073
## Year2004 -0.0493 0.2646 -0.19 0.8530
## Year2005 -0.1840 0.1936 -0.95 0.3458
## Year2006 -0.3449 0.1876 -1.84 0.0711 .
## Year2007 -0.0233 0.2297 -0.10 0.9195
```

```

## Year2008          -0.2790      0.3055   -0.91    0.3649
## Year2009          -0.6360      0.2240   -2.84    0.0062 **
## Year2010          -0.1117      0.2380   -0.47    0.6405
## Year2011           0.0900      0.2969    0.30    0.7629
## Year2012          -0.2971      0.2116   -1.40    0.1655
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.388
## Multiple R-squared:  0.19,   Adjusted R-squared:  -0.0297
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.349  0.875  0.960  0.903  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          1.32e-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: 76"
## [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]"
##
## 2006 2011
##    1    2
##
## 2006 2011
##    1    2
##
## 2006 2011
##    1    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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 3"
## [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
##    2    1    5    3    3    8    5    1    4    2    8    5    3    5    2
## 2011 2012
##    5    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    1    0    1    0    2    0    2    2    6    3    0    5    2
## 2011 2012
##    4    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    0    1    0    0    0    2    0    2    2    6    3    0    4    2
## 2011 2012
##    4    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 3"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male last 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"
##
GVIF Df GVIF^(1/(2*Df))

```



```

## FirstAuthorFemale  5.95  1          2.439
## LastAuthorFemale  14.01  1          3.743
## Year                55.14  9          1.250

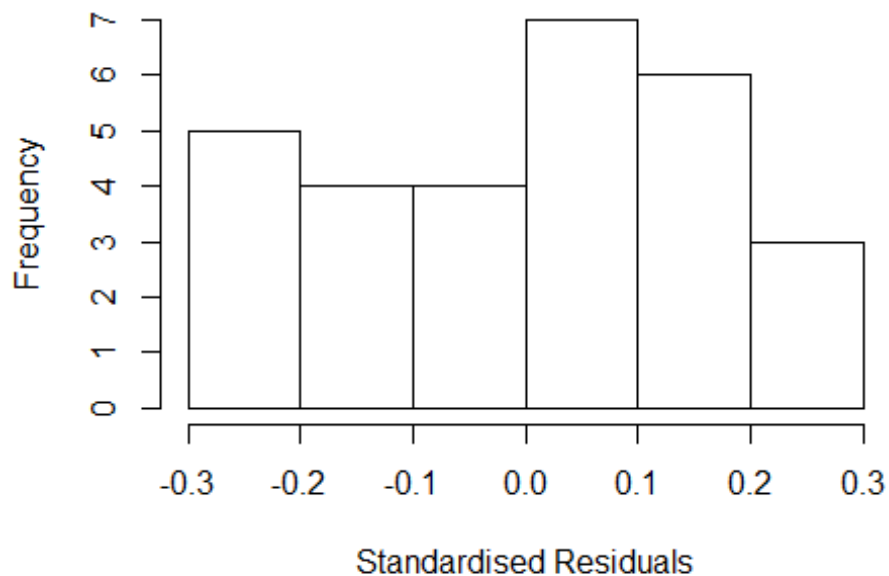
## [1] "List of  0 outliers with residuals above  2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields      residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.2840 -0.1699  0.0208  0.1558  0.2944
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    7.00e-01   7.02e-09  9.97e+07  < 2e-16 ***
## FirstAuthorFemale1 -2.74e-02   8.79e-02 -3.10e-01   0.7589
## LastAuthorFemale1 -1.16e-01   1.08e-01 -1.08e+00   0.2963
## Year2002         4.44e-01   1.15e-01  3.88e+00   0.0012 **
## Year2004         5.89e-01   6.12e-02  9.62e+00   2.7e-08 ***
## Year2005         4.13e-01   3.03e-01  1.36e+00   0.1912
## Year2006         7.25e-01   1.10e-01  6.57e+00   4.8e-06 ***
## Year2007         6.93e-01   9.14e-02  7.59e+00   7.4e-07 ***
## Year2009         6.40e-01   1.09e-01  5.88e+00   1.8e-05 ***
## Year2010         3.27e-01   1.47e-01  2.23e+00   0.0397 *
## Year2011         6.15e-01   1.11e-01  5.55e+00   3.5e-05 ***
## Year2012         4.11e-01   1.17e-01  3.51e+00   0.0027 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.215
## Multiple R-squared:  0.501, Adjusted R-squared:  0.177
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 26 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.837  0.915  0.946  0.939  0.989  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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"

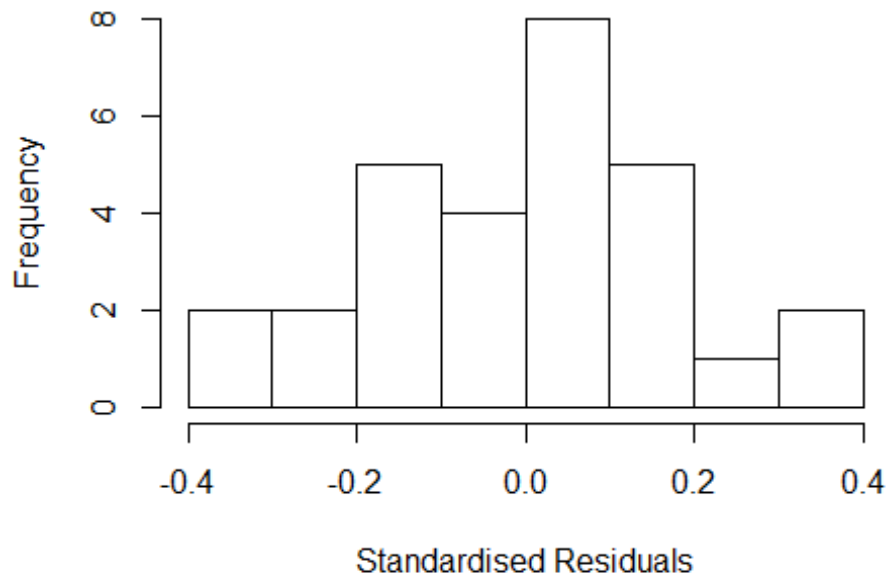
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```

Residuals from first and last author



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 15.92 1 3.991
## Year 15.92 9 1.166
```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.3867 -0.1099 0.0343 0.1035 0.3441
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7000 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.0401 0.0819 -0.49 0.63017
## Year2002 0.3280 0.0386 8.50 1.0e-07 ***
## Year2004 0.5371 0.0425 12.64 2.2e-10 ***
## Year2005 0.3611 0.4597 0.79 0.44243
## Year2006 0.7239 0.1224 5.91 1.4e-05 ***
## Year2007 0.6565 0.0653 10.05 8.3e-09 ***
## Year2009 0.6214 0.1278 4.86 0.00013 ***
## Year2010 0.2756 0.0878 3.14 0.00567 **
## Year2011 0.5939 0.0965 6.15 8.3e-06 ***
## Year2012 0.3775 0.1004 3.76 0.00143 **
## ---
```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.227
## Multiple R-squared:  0.475, Adjusted R-squared:  0.183
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 3 weights are ~= 1. The remaining 26 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.753  0.925  0.980  0.940  0.990  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           3.45e-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 15.86 1           3.982
## Year           15.86 9           1.166

## [1] "List of 0 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
## -0.2820 -0.1683  0.0167  0.1492  0.3055
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    7.00e-01   4.96e-09  1.41e+08 < 2e-16 ***
## LastAuthorFemale1 -1.22e-01   9.91e-02 -1.23e+00  0.2332
## Year2002         4.50e-01   1.06e-01  4.23e+00  0.0005 ***
## Year2004         5.78e-01   5.26e-02  1.10e+01  2.1e-09 ***

```

```

## Year2005          4.02e-01    2.44e-01    1.65e+00    0.1163
## Year2006          7.01e-01    7.25e-02    9.67e+00    1.5e-08 ***
## Year2007          6.86e-01    1.03e-01    6.68e+00    2.9e-06 ***
## Year2009          6.28e-01    1.00e-01    6.26e+00    6.6e-06 ***
## Year2010          3.17e-01    1.56e-01    2.03e+00    0.0578 .
## Year2011          5.96e-01    9.68e-02    6.15e+00    8.3e-06 ***
## Year2012          3.94e-01    1.24e-01    3.17e+00    0.0053 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.244
## Multiple R-squared:  0.495, Adjusted R-squared:  0.215
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.862  0.925  0.953  0.948  0.983  0.996
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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: 29"
## [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 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    4    4    6    3    4    3   10   10    8   13   15   15   17   23
## 2012
##    17
##
## 1996 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    3    2    1    1    2    2    5    3    6   10    8   13   11   19

```

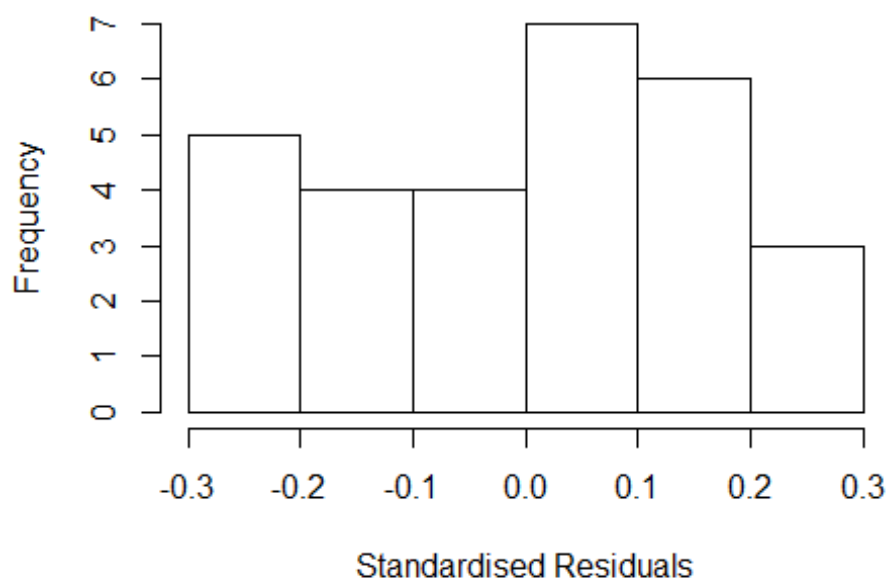
```
## 2012
## 16
##
## 1996 1997 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 0 2 2 1 1 2 2 5 3 6 8 8 12 11 19
## 2012
## 14
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.6672686083966"
## [1] "Male first author team size 2018 geometric mean: 4.37951913988789"

## 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 = 3.5, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male last author team size 2018 geometric mean: 4.21286593061052"

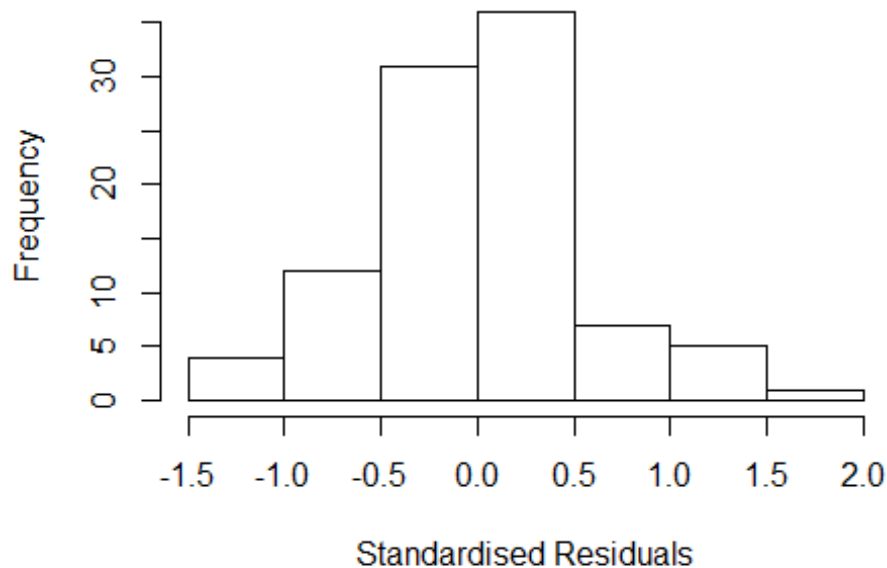
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

Residuals from last author



```
##
## 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"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.390e+00  1      1.841e+00
## LastAuthorFemale  2.402e+14  1      1.550e+07
## UniqueAuthors    3.130e+16  4      1.153e+02
## Year              2.326e+17 14      4.171e+00
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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:
```

```

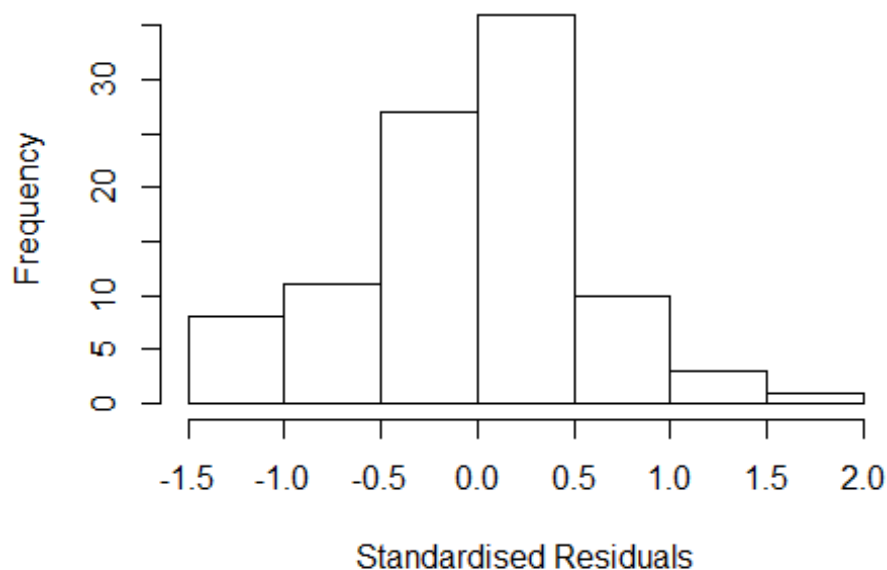
##      Min      1Q  Median      3Q      Max
## -1.2023 -0.3766  0.0159  0.3326  1.5426
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.2828     0.3517   0.80  0.4240
## FirstAuthorFemale1 -0.1188     0.1515  -0.78  0.4354
## LastAuthorFemale1 -0.0122     0.1774  -0.07  0.9455
## UniqueAuthors2      0.2652     0.2935   0.90  0.3690
## UniqueAuthors3      0.3119     0.2665   1.17  0.2456
## UniqueAuthors4      0.4500     0.2666   1.69  0.0956 .
## UniqueAuthors5      0.2205     0.4250   0.52  0.6054
## Year1998            0.4177     0.2998   1.39  0.1676
## Year1999            0.4542     0.3517   1.29  0.2005
## Year2000            0.6885     0.2903   2.37  0.0203 *
## Year2002           -0.1453     0.2902  -0.50  0.6180
## Year2003            0.1520     0.3130   0.49  0.6286
## Year2004            0.5402     0.2197   2.46  0.0163 *
## Year2005            0.5501     0.2785   1.98  0.0519 .
## Year2006            0.9931     0.3197   3.11  0.0027 **
## Year2007            0.5860     0.2854   2.05  0.0436 *
## Year2008            0.6260     0.3066   2.04  0.0447 *
## Year2009            0.7235     0.4157   1.74  0.0859 .
## Year2010            0.4695     0.3384   1.39  0.1695
## Year2011            0.5136     0.3034   1.69  0.0946 .
## Year2012            0.5107     0.3242   1.58  0.1194
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.552
## Multiple R-squared:  0.138, Adjusted R-squared:  -0.092
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 85 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.416  0.835  0.958  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          1.04e-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 2.406e+00 1 1.551e+00
## LastAuthorFemale 1.527e+15 1 3.908e+07
## Year 4.585e+15 14 3.625e+00
```

Residuals from first and last author



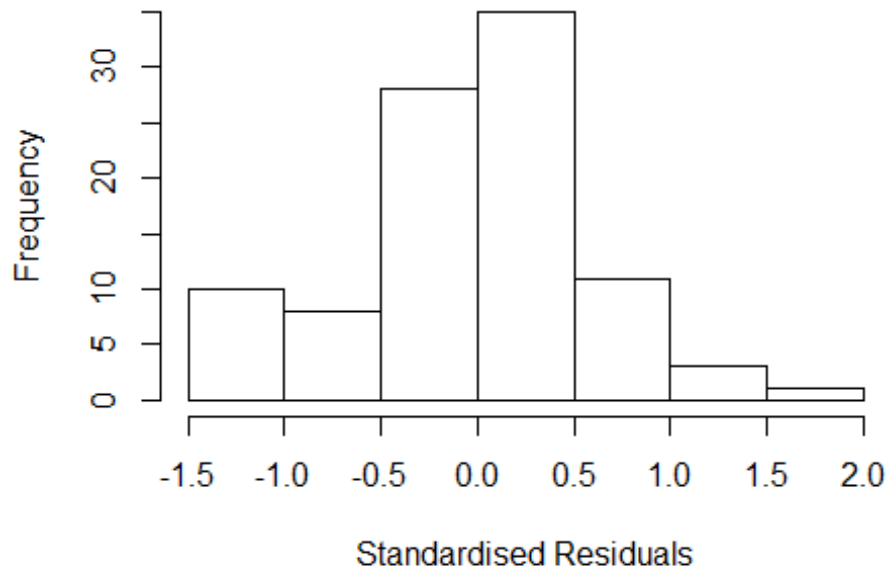
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1797 -0.3261 0.0337 0.3063 1.5554
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6180 0.0834 7.41 1.2e-10 ***
## FirstAuthorFemale1 -0.1011 0.1572 -0.64 0.52200
```

```

## LastAuthorFemale1 -0.0359 0.1820 -0.20 0.84400
## Year1998 0.4520 0.1575 2.87 0.00528 **
## Year1999 0.1190 0.0834 1.43 0.15737
## Year2000 0.6889 0.2002 3.44 0.00093 ***
## Year2002 -0.1920 0.1215 -1.58 0.11813
## Year2003 0.0850 0.1685 0.50 0.61527
## Year2004 0.5783 0.1405 4.11 9.4e-05 ***
## Year2005 0.4020 0.1659 2.42 0.01769 *
## Year2006 0.8557 0.2337 3.66 0.00045 ***
## Year2007 0.5380 0.1600 3.36 0.00119 **
## Year2008 0.4972 0.2846 1.75 0.08448 .
## Year2009 0.6406 0.3320 1.93 0.05722 .
## Year2010 0.4415 0.2411 1.83 0.07090 .
## Year2011 0.4661 0.2017 2.31 0.02347 *
## Year2012 0.5148 0.2333 2.21 0.03022 *
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.575
## Multiple R-squared: 0.095, Adjusted R-squared: -0.0883
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 88 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.444 0.853 0.968 0.908 0.992 0.999
## Algorithmic parameters:
## tuning.chi bb tuning.psi refine.tol
## 1.55e+00 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-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 2.906 1 1.705
## Year 2.906 14 1.039

```

Residuals from first author



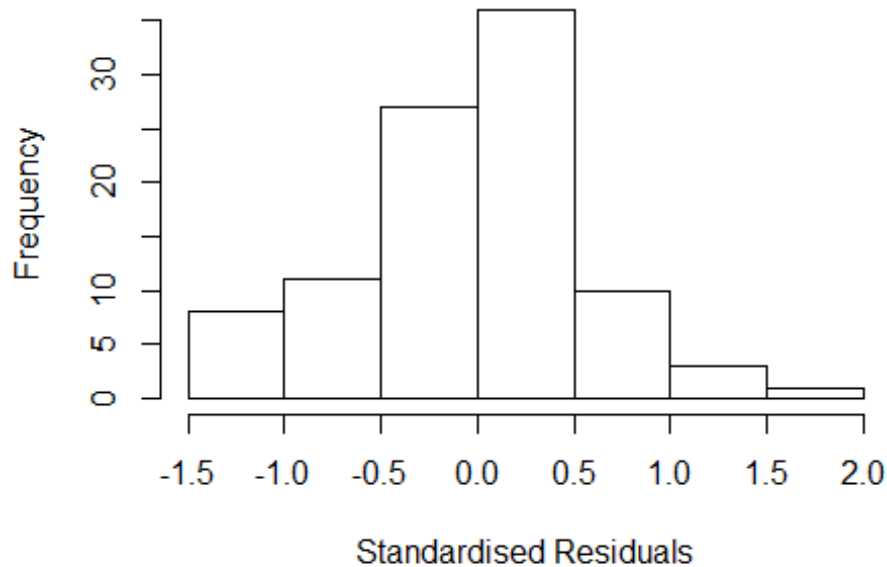
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1700 -0.3271 0.0413 0.3113 1.5602
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6180 0.0834 7.41 1.1e-10 ***
## FirstAuthorFemale1 -0.1047 0.1634 -0.64 0.52352
## Year1998 0.4340 0.1190 3.65 0.00047 ***
## Year1999 0.1190 0.0834 1.43 0.15746
## Year2000 0.6530 0.0834 7.83 1.7e-11 ***
## Year2002 -0.1920 0.1216 -1.58 0.11821
## Year2003 0.0689 0.1457 0.47 0.63770
## Year2004 0.5783 0.1406 4.11 9.4e-05 ***
## Year2005 0.3907 0.1679 2.33 0.02249 *
## Year2006 0.8460 0.2127 3.98 0.00015 ***
## Year2007 0.5313 0.1545 3.44 0.00093 ***
## Year2008 0.4977 0.2961 1.68 0.09664 .
```

```

## Year2009          0.6358      0.3446      1.84  0.06874 .
## Year2010          0.4411      0.2416      1.83  0.07162 .
## Year2011          0.4518      0.1760      2.57  0.01214 *
## Year2012          0.5051      0.2232      2.26  0.02631 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.563
## Multiple R-squared:  0.0953, Adjusted R-squared:  -0.0743
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.422  0.847   0.966   0.904   0.993   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 9e+13 1      9.487e+06
## Year              9e+13 14      3.150e+00

```

Residuals from last author



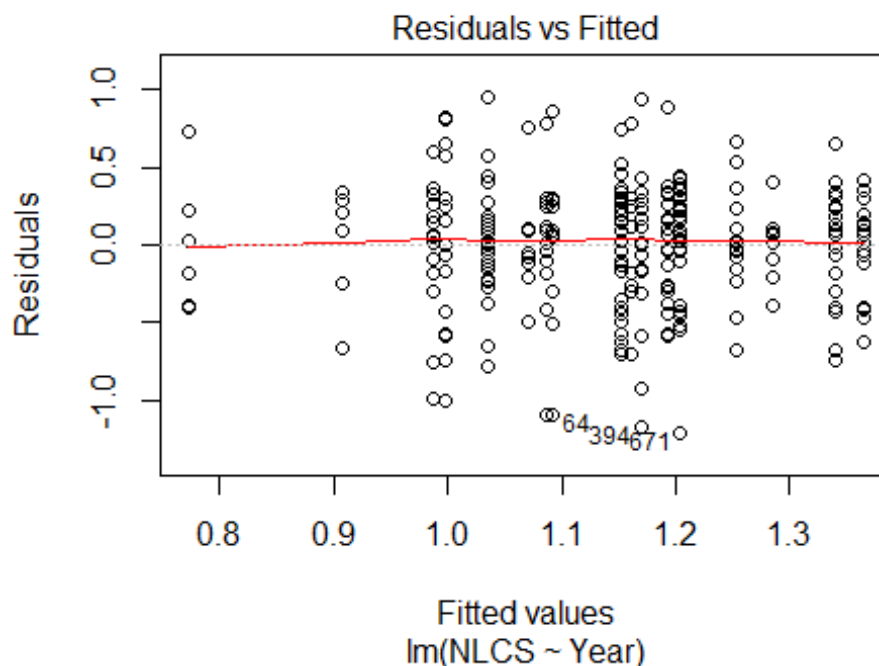
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2239 -0.2996 0.0771 0.2855 1.5901
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6180 0.0834 7.41 1.1e-10 ***
## LastAuthorFemale1 -0.0514 0.1915 -0.27 0.78897
## Year1998 0.4597 0.1639 2.81 0.00630 **
## Year1999 0.1190 0.0834 1.43 0.15746
## Year2000 0.7044 0.2089 3.37 0.00115 **
## Year2002 -0.1920 0.1216 -1.58 0.11821
## Year2003 0.0422 0.1658 0.25 0.79965
## Year2004 0.5783 0.1406 4.11 9.4e-05 ***
## Year2005 0.4070 0.1636 2.49 0.01492 *
## Year2006 0.8258 0.2157 3.83 0.00026 ***
## Year2007 0.4936 0.1430 3.45 0.00089 ***
## Year2008 0.4749 0.3044 1.56 0.12273
```

```

## Year2009          0.6059      0.3558      1.70  0.09246 .
## Year2010          0.4059      0.2317      1.75  0.08362 .
## Year2011          0.4376      0.1892      2.31  0.02328 *
## Year2012          0.4734      0.2245      2.11  0.03806 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.563
## Multiple R-squared:  0.0919, Adjusted R-squared:  -0.0784
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.405  0.862  0.966  0.904  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      1.04e-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: 96"
## [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
##   28   28   19   23   19   29   25   28   25   32   39   29   46   25   37
## 2011 2012
##   40   41
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8   10    6    6    8    8   12   18   16   18   22   18   30   15   20
## 2011 2012
##   30   28
##

```

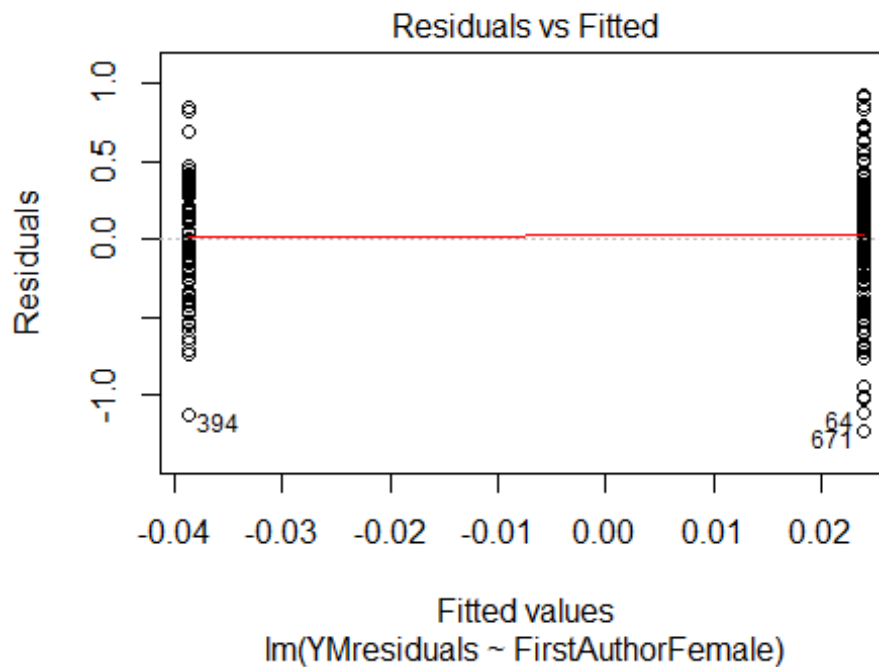
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7   10    6    5    5    8    8   14   15   15   20   16   28   14   19
## 2011 2012
##   29   26
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 17, df = 16, p-value = 0.4
```



```
##
## 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: 5.96105235797432"
## [1] "Male first author team size 2018 geometric mean: 3.75099400918396"
##
## 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 = 50, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.74582050814075"
## [1] "Male last author team size 2018 geometric mean: 4.04613838365249"

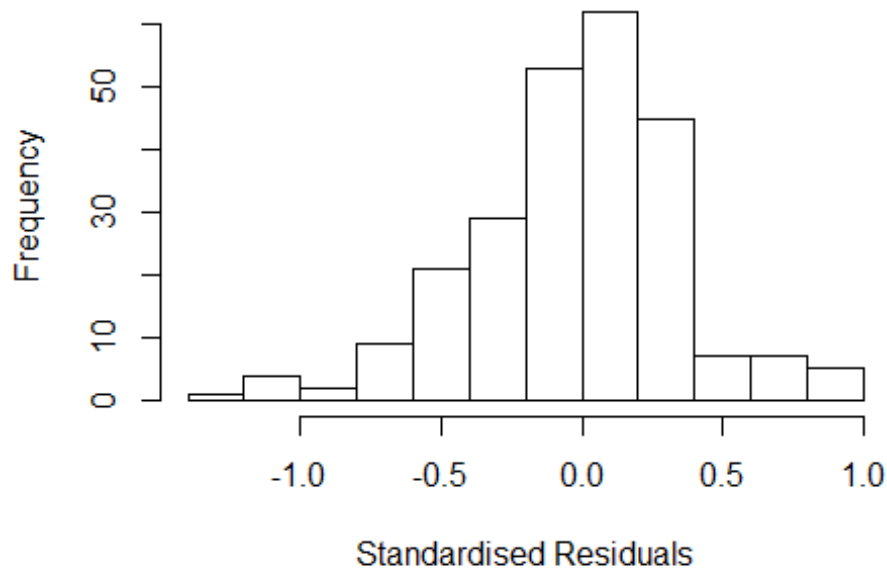
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 48, 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	1.419	1	1.191
LastAuthorFemale	1.797	1	1.341
UniqueAuthors	15.086	4	1.404
Year	17.796	16	1.094

Residuals from first and last author and team size



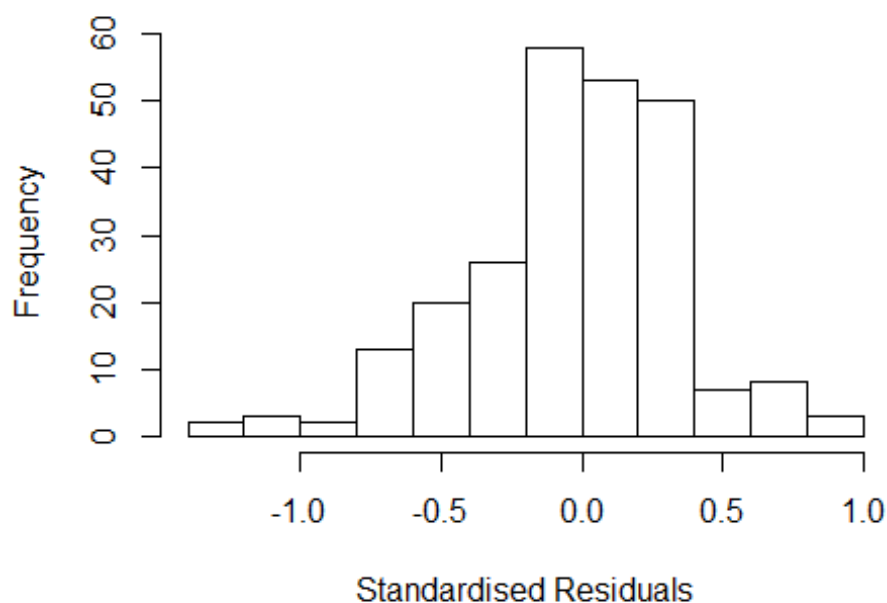
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3283 -0.2073  0.0183  0.2047  1.0000
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0982    0.2120   5.18   5e-07 ***
## FirstAuthorFemale1 -0.1091    0.0516  -2.11   0.0355 *
## LastAuthorFemale1  0.1036    0.0577   1.80   0.0739 .
## UniqueAuthors2     0.0983    0.1999   0.49   0.6234
## UniqueAuthors3     0.1724    0.2045   0.84   0.4000
## UniqueAuthors4     0.1996    0.1805   1.11   0.2700
## UniqueAuthors5     0.2797    0.1770   1.58   0.1156
## Year1997          -0.1605    0.2132  -0.75   0.4523
## Year1998          -0.4952    0.1857  -2.67   0.0082 **
## Year1999          -0.3430    0.1703  -2.01   0.0452 *
```

```

## Year2000          -0.0196      0.2415   -0.08   0.9355
## Year2001          -0.3190      0.1588   -2.01   0.0458 *
## Year2002          -0.3698      0.1832   -2.02   0.0448 *
## Year2003          -0.2751      0.1608   -1.71   0.0886 .
## Year2004          -0.0512      0.1580   -0.32   0.7461
## Year2005           0.1507      0.1372    1.10   0.2735
## Year2006          -0.1135      0.1308   -0.87   0.3862
## Year2007           0.0284      0.1653    0.17   0.8636
## Year2008          -0.2787      0.1329   -2.10   0.0371 *
## Year2009          -0.3348      0.2283   -1.47   0.1440
## Year2010           0.0463      0.1441    0.32   0.7484
## Year2011          -0.1234      0.1424   -0.87   0.3873
## Year2012          -0.0496      0.1339   -0.37   0.7112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.319
## Multiple R-squared:  0.235, Adjusted R-squared:  0.159
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 231 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0445 0.8700 0.9590 0.8840 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-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 1.188 1 1.090
## LastAuthorFemale 1.410 1 1.187
## Year 1.354 16 1.010

```

Residuals from first and last author



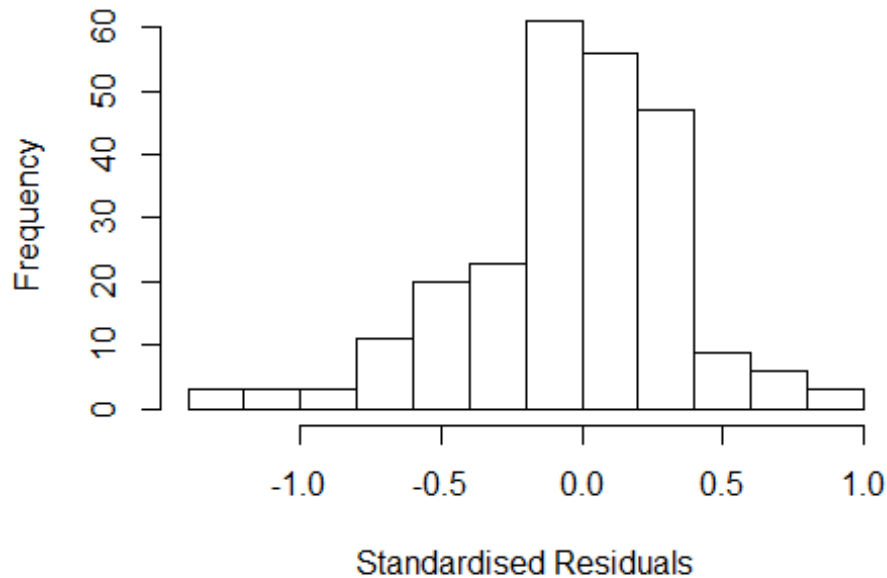
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.30333 -0.22511 -0.00136 0.22050 0.95430
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.30167 0.09103 14.30 <2e-16 ***
## FirstAuthorFemale1 -0.11661 0.05064 -2.30 0.0222 *
## LastAuthorFemale1 0.09783 0.05391 1.81 0.0709 .
## Year1997 -0.10690 0.15656 -0.68 0.4954
## Year1998 -0.55784 0.18632 -2.99 0.0031 **
## Year1999 -0.26648 0.17100 -1.56 0.1206
## Year2000 -0.13778 0.18309 -0.75 0.4525
## Year2001 -0.29308 0.14240 -2.06 0.0407 *
## Year2002 -0.32356 0.16513 -1.96 0.0513 .
## Year2003 -0.23125 0.13902 -1.66 0.0976 .
## Year2004 -0.02032 0.11989 -0.17 0.8656
## Year2005 0.17521 0.11148 1.57 0.1174
```

```

## Year2006      -0.09510    0.11234   -0.85    0.3982
## Year2007      0.02066    0.13741    0.15    0.8806
## Year2008     -0.26317    0.10725   -2.45    0.0149 *
## Year2009     -0.32736    0.21368   -1.53    0.1269
## Year2010      0.09155    0.12497    0.73    0.4646
## Year2011     -0.09157    0.12295   -0.74    0.4572
## Year2012      0.00166    0.11208    0.01    0.9882
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.32
## Multiple R-squared:  0.201, Adjusted R-squared:  0.137
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 224 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0604 0.8600 0.9470 0.8780 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-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 1.096 1      1.047
## Year      1.096 16      1.003

```

Residuals from first author



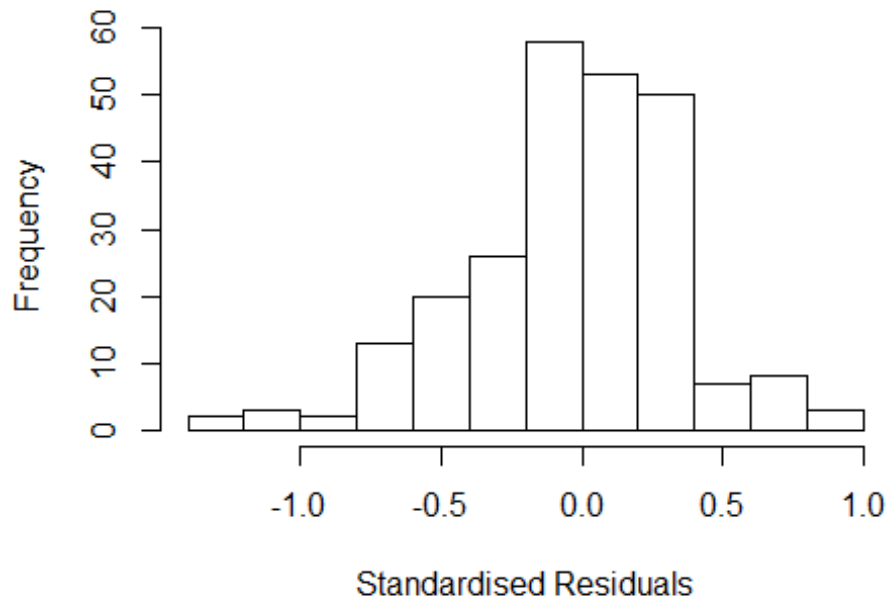
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.32033 -0.20817 -0.00195 0.21378 0.91472
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.31461 0.09320 14.11 <2e-16 ***
## FirstAuthorFemale1 -0.09876 0.04924 -2.01 0.0461 *
## Year1997 -0.11129 0.16688 -0.67 0.5055
## Year1998 -0.57508 0.18853 -3.05 0.0026 **
## Year1999 -0.25938 0.17923 -1.45 0.1492
## Year2000 -0.13800 0.18012 -0.77 0.4444
## Year2001 -0.26532 0.14403 -1.84 0.0668 .
## Year2002 -0.27707 0.14481 -1.91 0.0570 .
## Year2003 -0.22425 0.13302 -1.69 0.0932 .
## Year2004 0.00631 0.11896 0.05 0.9577
## Year2005 0.18660 0.11403 1.64 0.1031
## Year2006 -0.08965 0.11523 -0.78 0.4374
```

```

## Year2007          0.02053    0.13500    0.15    0.8793
## Year2008         -0.25628    0.10935   -2.34    0.0200 *
## Year2009         -0.31857    0.22958   -1.39    0.1666
## Year2010          0.09034    0.12671    0.71    0.4766
## Year2011         -0.09742    0.12687   -0.77    0.4434
## Year2012          0.00572    0.11517    0.05    0.9605
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.319
## Multiple R-squared:  0.19,   Adjusted R-squared:  0.129
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 230 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0476 0.8610 0.9530 0.8780 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2681 -0.2256 0.0249 0.1949 0.9102
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2871 0.0898 14.34 <2e-16 ***
## LastAuthorFemale1 0.0701 0.0505 1.39 0.1664
## Year1997 -0.1199 0.1628 -0.74 0.4623
## Year1998 -0.5671 0.1976 -2.87 0.0045 **
## Year1999 -0.2732 0.1649 -1.66 0.0991 .
## Year2000 -0.1997 0.2007 -1.00 0.3206
## Year2001 -0.2835 0.1345 -2.11 0.0361 *
## Year2002 -0.3428 0.1555 -2.20 0.0285 *
## Year2003 -0.2604 0.1422 -1.83 0.0683 .
## Year2004 -0.0391 0.1220 -0.32 0.7490
## Year2005 0.1382 0.1087 1.27 0.2049
## Year2006 -0.1142 0.1130 -1.01 0.3133
```

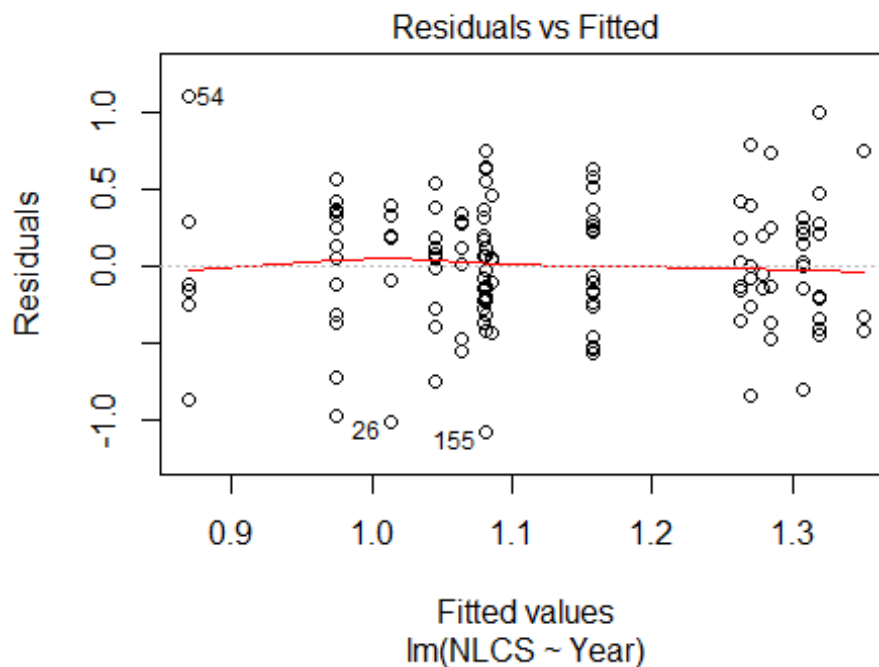
```

## Year2007          -0.0198      0.1354   -0.15   0.8836
## Year2008          -0.2880      0.1062   -2.71   0.0072 **
## Year2009          -0.3551      0.2316   -1.53   0.1266
## Year2010           0.0870      0.1229    0.71   0.4795
## Year2011          -0.1257      0.1255   -1.00   0.3179
## Year2012          -0.0190      0.1104   -0.17   0.8635
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.318
## Multiple R-squared:  0.181, Adjusted R-squared:  0.12
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 224 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.076  0.852  0.950  0.874  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      4.08e-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: 245"
## [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
##    5    9    7   17   13    8    7    8    9    8   10   12   25   23   16
## 2011 2012
##   29   18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5    6   13    6    3    6    3    7    5    7    8   15   10    8
## 2011 2012

```



```
## 18 11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 5 5 13 3 3 5 3 6 5 6 6 11 9 7
## 2011 2012
## 16 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.5
```



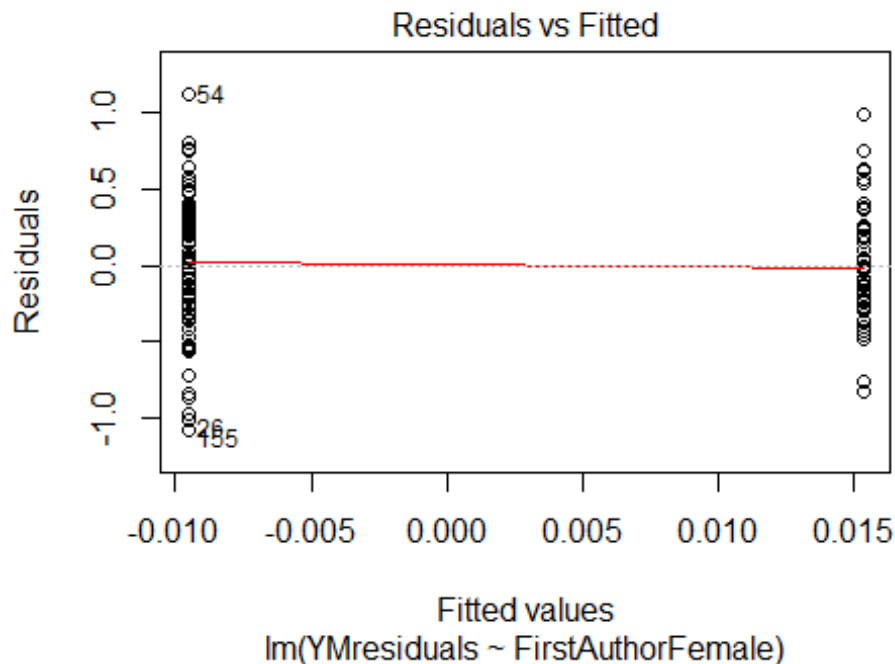
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.3, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 5.4400868459948"
## [1] "Male first author team size 2018 geometric mean: 4.48140474655716"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

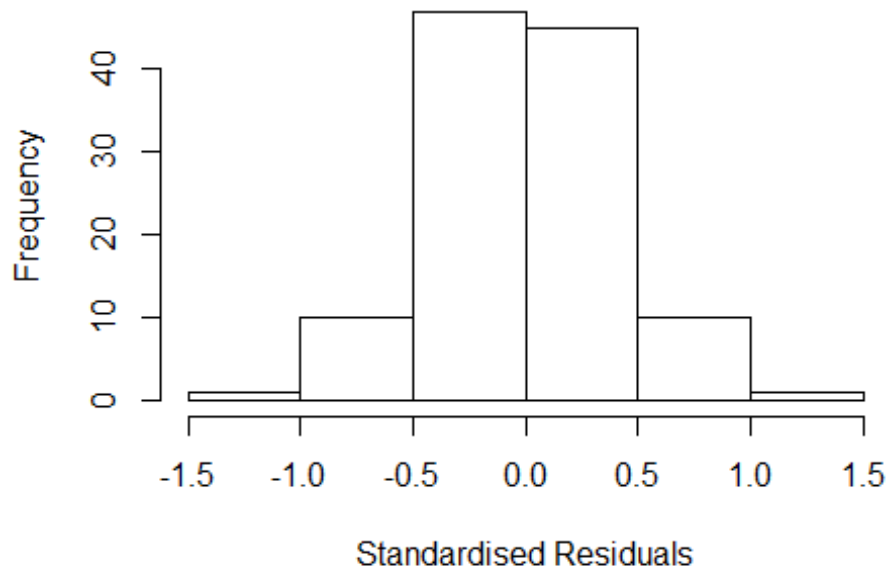
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.64398437900346"
## [1] "Male last author team size 2018 geometric mean: 5.7326567532262"

## 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 = 8.5, 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  5.922  1      2.434
## LastAuthorFemale  2.990  1      1.729
## UniqueAuthors    58.913  4      1.664
## Year             199.798 16      1.180
```

Residuals from first and last author and team size



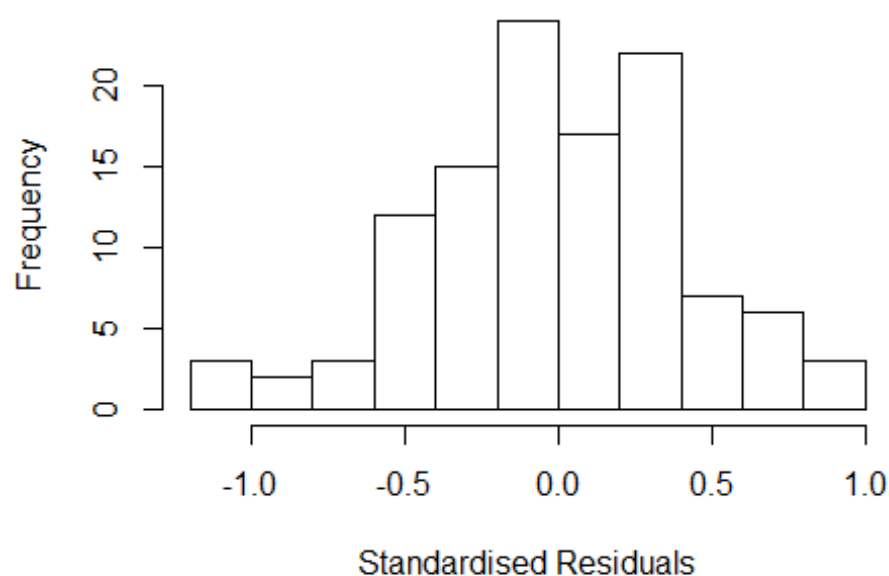
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.233820 -0.248397 -0.000298 0.261289 1.100818
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.66842 0.14759 4.53 1.8e-05 ***
## FirstAuthorFemale1 -0.08764 0.10161 -0.86 0.39070
## LastAuthorFemale1 -0.09557 0.10567 -0.90 0.36815
## UniqueAuthors2 0.46858 0.16381 2.86 0.00525 **
## UniqueAuthors3 0.33423 0.15447 2.16 0.03311 *
## UniqueAuthors4 0.52622 0.14094 3.73 0.00033 ***
## UniqueAuthors5 0.45813 0.14458 3.17 0.00209 **
## Year1997 0.25342 0.21773 1.16 0.24750
## Year1998 -0.00852 0.26345 -0.03 0.97429
## Year1999 0.09682 0.13855 0.70 0.48645
```

```

## Year2000      0.38232    0.38177    1.00  0.31926
## Year2001      0.58609    0.42280    1.39  0.16907
## Year2002      0.28963    0.29269    0.99  0.32503
## Year2003      0.25123    0.21662    1.16  0.24917
## Year2004      0.42775    0.18503    2.31  0.02305 *
## Year2005      0.06024    0.14824    0.41  0.68544
## Year2006      0.08092    0.15908    0.51  0.61224
## Year2007      0.30596    0.17144    1.78  0.07764 .
## Year2008      0.12801    0.23147    0.55  0.58160
## Year2009      0.12478    0.10756    1.16  0.24907
## Year2010      0.30417    0.25497    1.19  0.23598
## Year2011      0.12648    0.18638    0.68  0.49909
## Year2012      0.12695    0.16226    0.78  0.43601
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.399
## Multiple R-squared:  0.202, Adjusted R-squared:  0.00856
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 106 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.317  0.882  0.959  0.912  0.981  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           8.77e-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 5.381 1           2.320
## LastAuthorFemale  2.657 1           1.630
## Year              11.349 16           1.079

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.1108 -0.3090 -0.0339 0.2923 0.9891
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.051733 0.103997 10.11 <2e-16 ***
## FirstAuthorFemale1 0.055267 0.103997 0.53 0.596
## LastAuthorFemale1 0.000466 0.115324 0.00 0.997
## Year1997 0.182787 0.240401 0.76 0.449
## Year1998 -0.016924 0.219070 -0.08 0.939
## Year1999 -0.040649 0.143985 -0.28 0.778
## Year2000 0.144544 0.508169 0.28 0.777
## Year2001 0.206768 0.416602 0.50 0.621
## Year2002 0.173567 0.320507 0.54 0.589
## Year2003 0.207588 0.111207 1.87 0.065 .
## Year2004 0.195635 0.140215 1.40 0.166
## Year2005 0.010760 0.156537 0.07 0.945
```

```

## Year2006          0.020316    0.183371    0.11    0.912
## Year2007          0.233295    0.180804    1.29    0.200
## Year2008          0.059027    0.209665    0.28    0.779
## Year2009          0.022377    0.089205    0.25    0.802
## Year2010          0.223903    0.264883    0.85    0.400
## Year2011          0.017527    0.134889    0.13    0.897
## Year2012          -0.024791    0.133857   -0.19    0.853
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.0562, Adjusted R-squared:  -0.123
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 106 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.430  0.883  0.944  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      8.77e-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 5.237 1      2.288
## Year              5.237 16      1.053
##
## [1] "List of 0 outliers 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

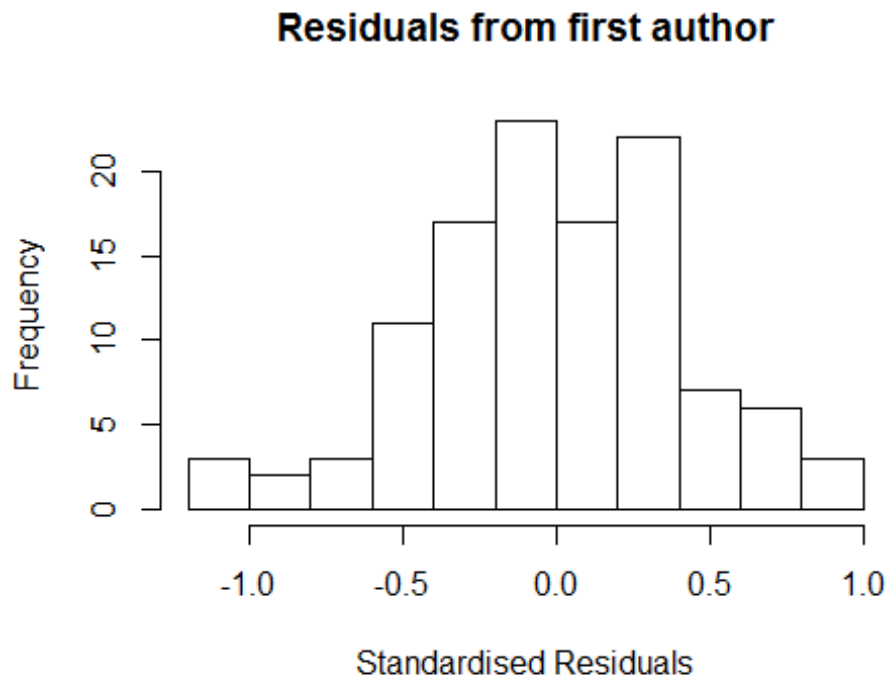
```

```

## -1.1100 -0.3102 -0.0333 0.2931 0.9863
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.0514    0.1031   10.20 <2e-16 ***
## FirstAuthorFemale1 0.0556    0.1031    0.54 0.591
## Year1997        0.1845    0.2319    0.80 0.428
## Year1998       -0.0199    0.2179   -0.09 0.927
## Year1999       -0.0420    0.1439   -0.29 0.771
## Year2000        0.1464    0.4935    0.30 0.767
## Year2001        0.2109    0.4135    0.51 0.611
## Year2002        0.1742    0.3207    0.54 0.588
## Year2003        0.2080    0.1022    2.03 0.045 *
## Year2004        0.1959    0.1395    1.40 0.163
## Year2005        0.0110    0.1561    0.07 0.944
## Year2006        0.0200    0.1811    0.11 0.912
## Year2007        0.2324    0.1654    1.40 0.163
## Year2008        0.0585    0.2074    0.28 0.778
## Year2009        0.0225    0.0890    0.25 0.801
## Year2010        0.2267    0.2361    0.96 0.339
## Year2011        0.0183    0.1152    0.16 0.874
## Year2012       -0.0250    0.1385   -0.18 0.857
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared: 0.0567, Adjusted R-squared: -0.11
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 106 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.443 0.886 0.945 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          8.77e-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"

```

```
## 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 2.517  1          1.587
## Year            2.517 16          1.029

## [1] "List of  0 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
## -1.12109 -0.30539 -0.00888  0.29221  1.00375
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.11e+00  0.00e+00    Inf <2e-16 ***
## LastAuthorFemale1  5.16e-04  1.11e-01   0.00  0.996
## Year1997         1.44e-01  2.36e-01   0.61  0.543
## Year1998        -7.88e-02  2.06e-01  -0.38  0.702
## Year1999        -8.36e-02  1.40e-01  -0.60  0.553
```



```

## Year2000      1.07e-01  4.45e-01  0.24  0.810
## Year2001      1.59e-01  3.91e-01  0.41  0.685
## Year2002      1.71e-01  3.17e-01  0.54  0.592
## Year2003      1.70e-01  9.74e-02  1.75  0.084 .
## Year2004      1.48e-01  1.06e-01  1.39  0.167
## Year2005     -2.29e-02  1.33e-01  -0.17  0.864
## Year2006     -2.47e-02  1.74e-01  -0.14  0.888
## Year2007      2.10e-01  1.70e-01  1.24  0.218
## Year2008      1.41e-02  1.93e-01  0.07  0.942
## Year2009      7.36e-05  8.65e-02  0.00  0.999
## Year2010      2.09e-01  2.47e-01  0.85  0.399
## Year2011     -9.57e-03  1.25e-01  -0.08  0.939
## Year2012     -5.52e-02  1.13e-01  -0.49  0.627
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.0541, Adjusted R-squared:  -0.113
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 105 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.447  0.887  0.949  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      8.77e-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: 114"
## [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 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012

```

```

##      1      2      1      4     10      4      2      5      4      4      8      4      6      5
##
## 1996 1997 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012
##      0      2      0      1      7      3      0      4      3      4      7      4      5      4
##
## 1996 1997 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012
##      0      2      0      0      7      3      0      3      2      4      7      4      4      3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1"
## [1] "Male first author team size 2018 geometric mean: 2.34901007932325"

## 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 = 1.5, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.64575131106459"
## [1] "Male last author team size 2018 geometric mean: 1.8881750225898"

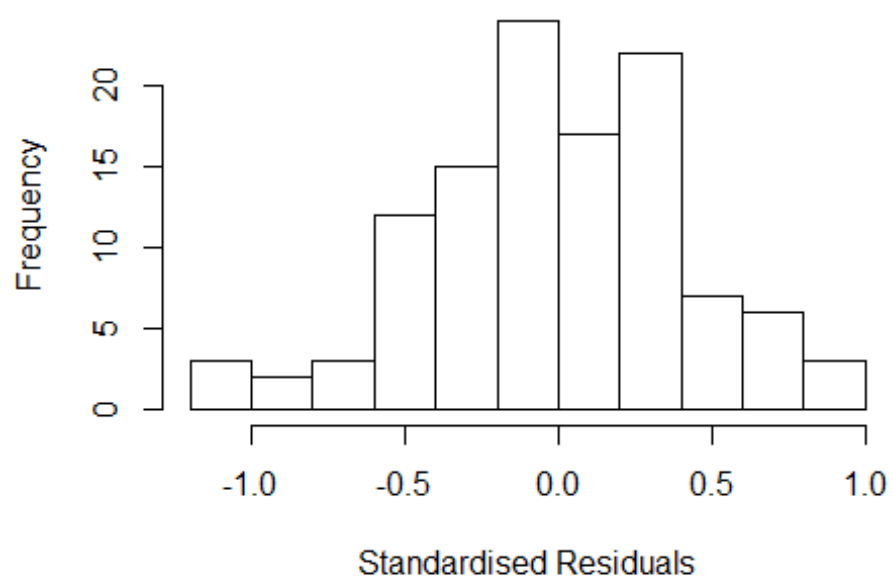
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5.5, 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"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate

```

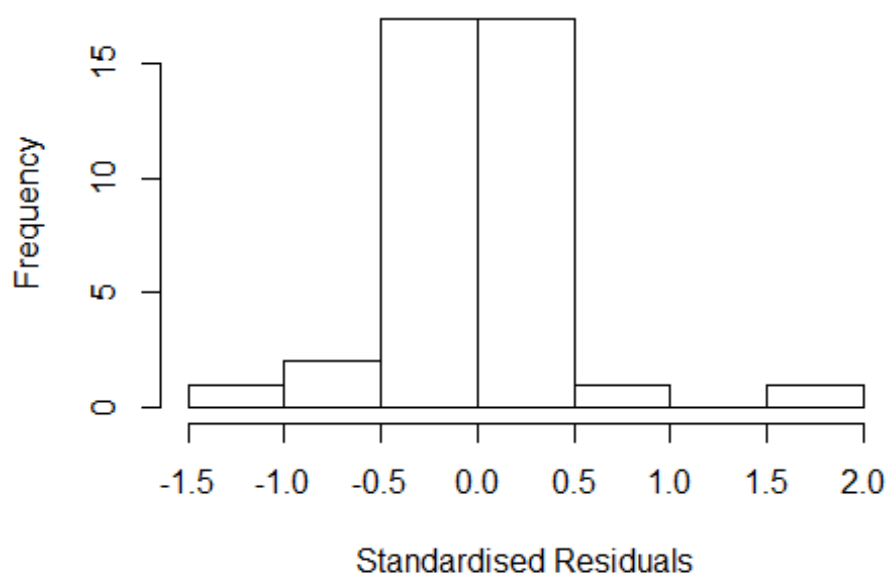
Residuals from last author



```
## [1] "Regression 2: First author gender, Last author gender, Year as factors"
```

##		GVIF	Df	$GVIF^{1/(2 \cdot Df)}$
##	FirstAuthorFemale	13.62	1	3.691
##	LastAuthorFemale	88.10	1	9.386
##	Year	1179.14	9	1.481

Residuals from first and last author



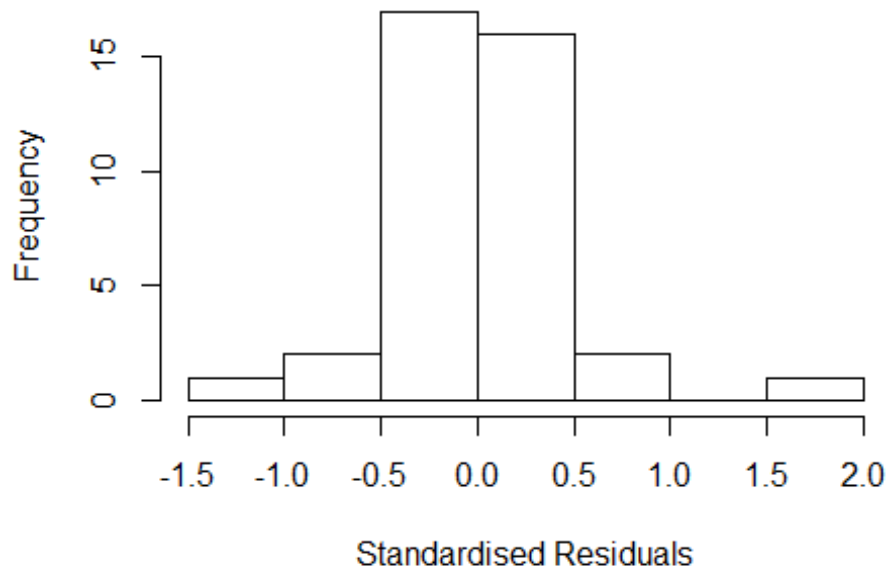
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20908 -0.12245 -0.00352 0.14986 1.57660
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.29500 0.00566 405.64 < 2e-16 ***
## FirstAuthorFemale1 0.36588 0.14172 2.58 0.01558 *
## LastAuthorFemale1 -0.00704 0.11474 -0.06 0.95156
## Year2002 -1.08592 0.24976 -4.35 0.00018 ***
## Year2003 -2.29148 0.05770 -39.71 < 2e-16 ***
## Year2005 -0.94145 0.16971 -5.55 7.0e-06 ***
## Year2006 -0.63242 0.11783 -5.37 1.1e-05 ***
## Year2007 -0.81117 0.11903 -6.81 2.6e-07 ***
## Year2008 -0.85039 0.17068 -4.98 3.2e-05 ***
## Year2009 -0.35517 0.16125 -2.20 0.03634 *
## Year2011 -1.09626 0.20872 -5.25 1.5e-05 ***
## Year2012 -0.45980 0.14619 -3.15 0.00401 **
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.28
## Multiple R-squared:  0.826, Adjusted R-squared:  0.755
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.0026);
## 5 weights are ~= 1. The remaining 33 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0223 0.8930 0.9670 0.8760 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           2.56e-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 16.51 1           4.063
## Year              16.51 9           1.169

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.14e+00 -1.33e-01  8.88e-16  1.59e-01  1.55e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.29500    0.00566  405.65 < 2e-16 ***
## FirstAuthorFemale1  0.39322    0.16594    2.37  0.0249 *
## Year2002     -1.15306    0.30670   -3.76  0.0008 ***
## Year2003     -2.29500    0.00566 -405.65 < 2e-16 ***
## Year2005     -0.97376    0.19072   -5.11  2.1e-05 ***
## Year2006     -0.64961    0.12551   -5.18  1.7e-05 ***
## Year2007     -0.83465    0.13970   -5.97  2.0e-06 ***
## Year2008     -0.84900    0.16434   -5.17  1.8e-05 ***
## Year2009     -0.38595    0.17956   -2.15  0.0404 *
## Year2011     -1.11539    0.18418   -6.06  1.6e-06 ***
## Year2012     -0.46820    0.14789   -3.17  0.0037 **
## ---
```

```

## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.305
## Multiple R-squared:  0.801, Adjusted R-squared:  0.73
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.0026);
## 4 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.132  0.931  0.974  0.896  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           2.56e-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 32.05  1           5.662
## Year              32.05  9           1.212
##
## [1] "List of  0 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
## -1.1838 -0.2098  0.0256  0.1995  2.0396
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    2.29500    0.00566  405.66 < 2e-16 ***
## LastAuthorFemale1 0.18726    0.18437    1.02  0.31848
## Year2002        -1.11118    0.39636   -2.80  0.00908 **

```

```

## Year2003          -2.38863      0.11419  -20.92  < 2e-16 ***
## Year2005          -0.70558      0.17813   -3.96  0.00047 ***
## Year2006          -0.54663      0.09412   -5.81  3.1e-06 ***
## Year2007          -0.63190      0.18365   -3.44  0.00184 **
## Year2008          -0.67305      0.16632   -4.05  0.00037 ***
## Year2009          -0.08659      0.15557   -0.56  0.58222
## Year2011          -1.08730      0.20364   -5.34  1.1e-05 ***
## Year2012          -0.32307      0.13197   -2.45  0.02089 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.749, Adjusted R-squared:  0.66
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## observation 10 is an outlier with |weight| = 0 ( < 0.0026);
## 9 weights are ~ = 1. The remaining 29 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.192  0.923  0.948   0.876   0.975   0.994
## Algorithmic parameters:
##           tuning.chi                bb           tuning.psi           refine.tol
##           1.55e+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-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: 39"
## [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
##    3    5    8    6   11   15    5    5    9    8    4    9   19    8   19
## 2011 2012
##   19   17
##

```



```

## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    2    3    1    1    6    4    3    4    2    3    6   13    6   13
## 2011 2012
##   14    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    2    3    1    1    4    4    3    3    2    2    6   12    5   11
## 2011 2012
##   11    9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.50156321196687"
## [1] "Male first author team size 2018 geometric mean: 6.54213262037718"

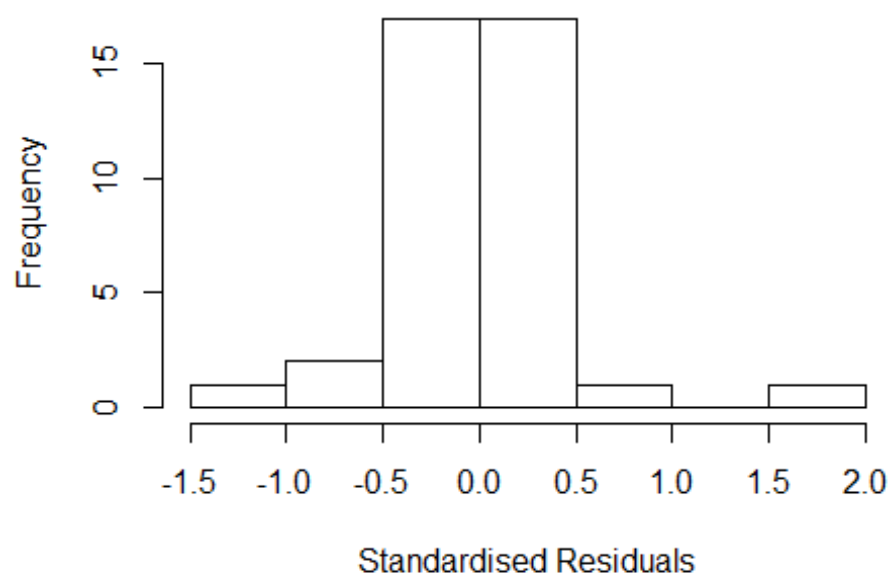
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.25732474567597"
## [1] "Male last author team size 2018 geometric mean: 5.65046900098953"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

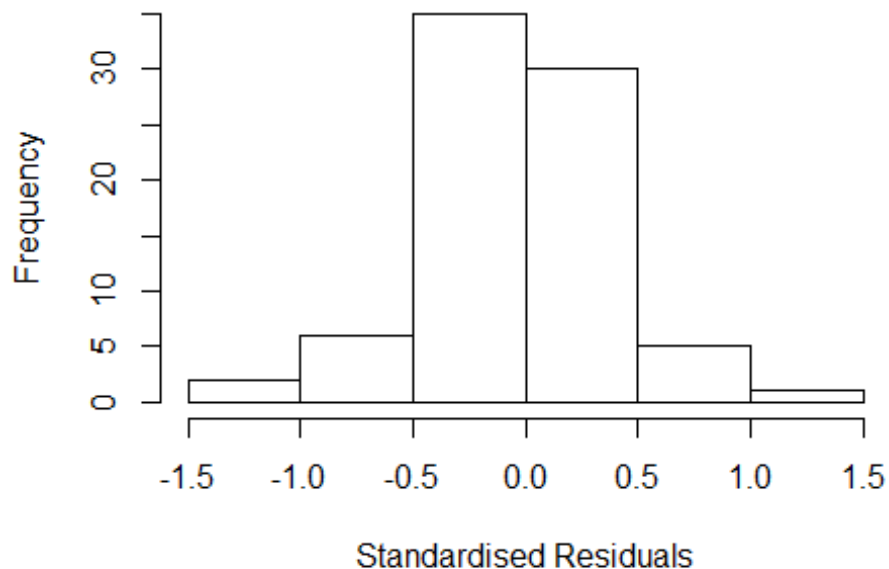
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8.5, 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 1.796e+01 1         4.238
## LastAuthorFemale  7.204e+00 1         2.684
## UniqueAuthors    -6.883e+18 4          NaN
## Year              -2.236e+19 15          NaN
```

Residuals from first and last author and team size



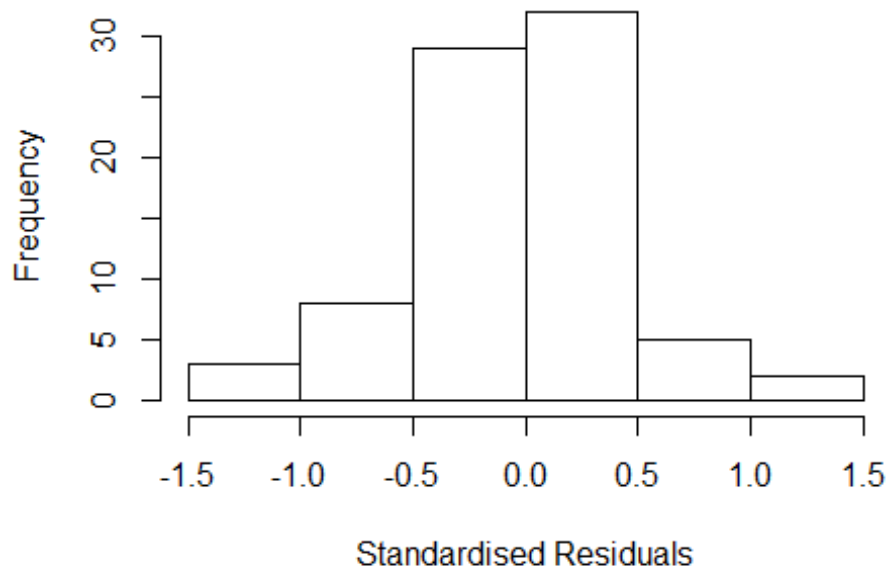
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.28110 -0.28113 -0.00635 0.29125 1.17521
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.51871 0.26936 9.35 4.2e-13 ***
## FirstAuthorFemale1 0.08295 0.13620 0.61 0.54492
## LastAuthorFemale1 -0.25012 0.14759 -1.69 0.09559 .
## UniqueAuthors2 -0.33778 0.31428 -1.07 0.28700
## UniqueAuthors3 -0.45763 0.25749 -1.78 0.08085 .
## UniqueAuthors4 -0.00157 0.30059 -0.01 0.99586
## UniqueAuthors5 -0.00233 0.27264 -0.01 0.99322
## Year1998 -0.96172 0.16115 -5.97 1.6e-07 ***
## Year1999 -1.17593 0.10921 -10.77 2.3e-15 ***
## Year2000 -0.43007 0.10921 -3.94 0.00023 ***
```

```

## Year2001      -1.11616    0.23112   -4.83  1.1e-05 ***
## Year2002      -0.91123    0.32702   -2.79  0.00723 **
## Year2003      -1.26904    0.14951   -8.49  1.1e-11 ***
## Year2004      -0.78327    0.72037   -1.09  0.28147
## Year2005      -1.18300    0.24276   -4.87  9.1e-06 ***
## Year2006      -0.81565    0.20991   -3.89  0.00027 ***
## Year2007      -1.39432    0.30435   -4.58  2.6e-05 ***
## Year2008      -1.02533    0.17027   -6.02  1.3e-07 ***
## Year2009      -1.18202    0.25313   -4.67  1.9e-05 ***
## Year2010      -0.93559    0.26879   -3.48  0.00097 ***
## Year2011      -1.25994    0.14989   -8.41  1.5e-11 ***
## Year2012      -1.20567    0.22413   -5.38  1.5e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.507
## Multiple R-squared:  0.304, Adjusted R-squared:  0.0471
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.503  0.917  0.965  0.927  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      1.27e-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 4.468 1 2.114
## LastAuthorFemale 3.531 1 1.879
## Year 11.020 15 1.083

```

Residuals from first and last author



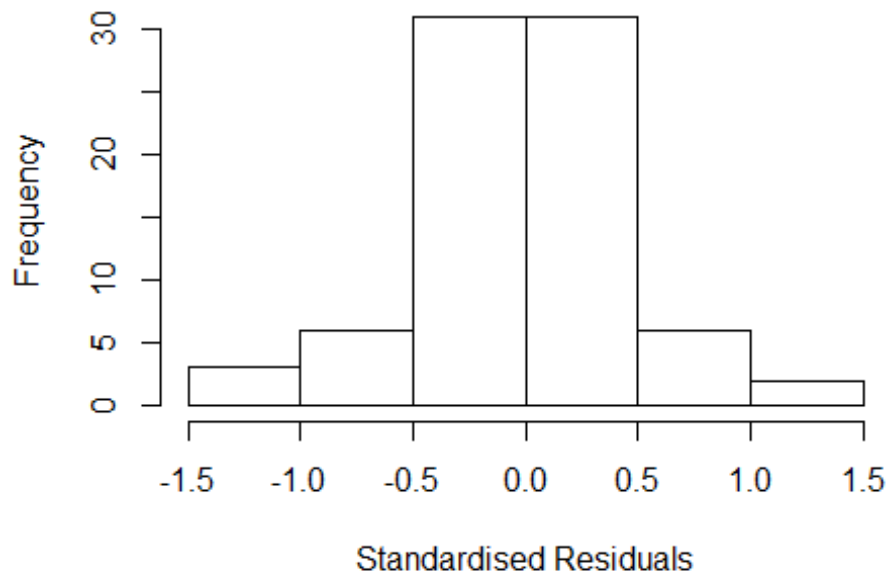
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.25e+00 -2.82e-01 5.55e-16 2.90e-01 1.22e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.1210 0.0741 28.63 < 2e-16 ***
## FirstAuthorFemale1 0.1215 0.1227 0.99 0.32592
## LastAuthorFemale1 -0.1989 0.1357 -1.47 0.14787
## Year1998 -0.7505 0.1742 -4.31 6.1e-05 ***
## Year1999 -1.1160 0.0741 -15.06 < 2e-16 ***
## Year2000 -0.4900 0.0741 -6.61 1.1e-08 ***
## Year2001 -0.8267 0.3000 -2.76 0.00771 **
## Year2002 -0.6680 0.3893 -1.72 0.09127 .
## Year2003 -1.0354 0.1540 -6.72 6.9e-09 ***
## Year2004 -0.4173 0.5968 -0.70 0.48706
## Year2005 -1.1830 0.2037 -5.81 2.4e-07 ***
## Year2006 -0.5880 0.3456 -1.70 0.09397 .
```

```

## Year2007          -1.0284      0.2938   -3.50  0.00088 ***
## Year2008          -0.7310      0.1524   -4.80  1.1e-05 ***
## Year2009          -0.8269      0.2186   -3.78  0.00036 ***
## Year2010          -0.5865      0.2263   -2.59  0.01194 *
## Year2011          -1.0630      0.1399   -7.60  2.2e-10 ***
## Year2012          -0.9257      0.1968   -4.70  1.5e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.514
## Multiple R-squared:  0.227, Adjusted R-squared:  0.0121
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.532  0.919  0.970  0.926  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      1.27e-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 4.41 1      2.100
## Year              4.41 15      1.051

```

Residuals from first author



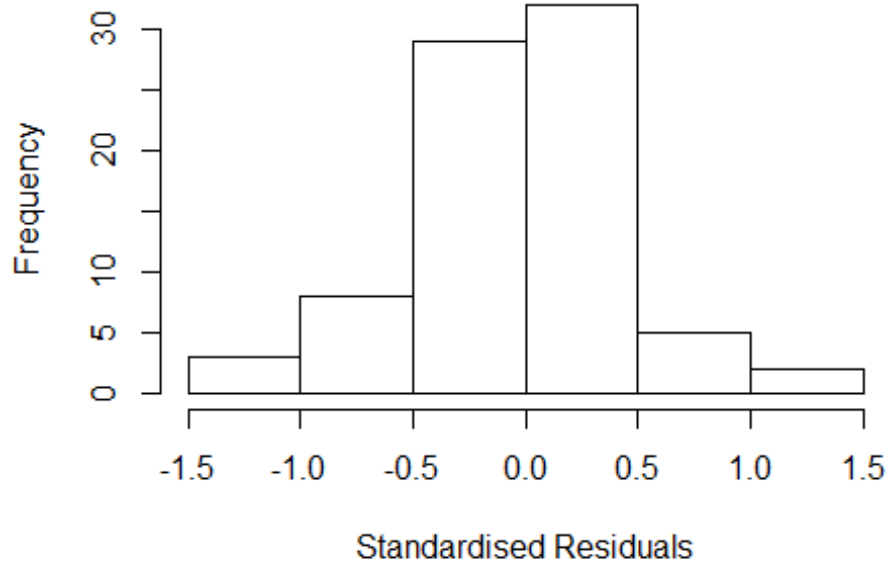
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.27e+00 -2.56e-01  2.22e-16  3.13e-01  1.30e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      2.1210     0.0741   28.62 < 2e-16 ***
## FirstAuthorFemale1  0.0894     0.1250    0.71  0.47754
## Year1998         -0.8844     0.1020   -8.67  2.8e-12 ***
## Year1999         -1.1160     0.0741  -15.06 < 2e-16 ***
## Year2000         -0.4900     0.0741   -6.61  1.0e-08 ***
## Year2001         -0.8165     0.3047   -2.68  0.00943 **
## Year2002         -0.7489     0.3478   -2.15  0.03522 *
## Year2003         -1.0243     0.1571   -6.52  1.4e-08 ***
## Year2004         -0.3941     0.6073   -0.65  0.51873
## Year2005         -1.1830     0.2038   -5.80  2.4e-07 ***
## Year2006         -0.5880     0.3464   -1.70  0.09463 .
## Year2007         -1.0511     0.2737   -3.84  0.00029 ***
```

```

## Year2008          -0.7818      0.1522    -5.14   3.0e-06 ***
## Year2009          -0.9017      0.1943    -4.64   1.8e-05 ***
## Year2010          -0.6677      0.2133    -3.13   0.00266 **
## Year2011          -1.0922      0.1467    -7.44   3.7e-10 ***
## Year2012          -0.9600      0.2235    -4.30   6.3e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.206, Adjusted R-squared:  0.000485
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 71 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.493  0.916  0.966  0.921  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      1.27e-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 3.117 1          1.765
## Year            3.117 15          1.039

```


Residuals from last author



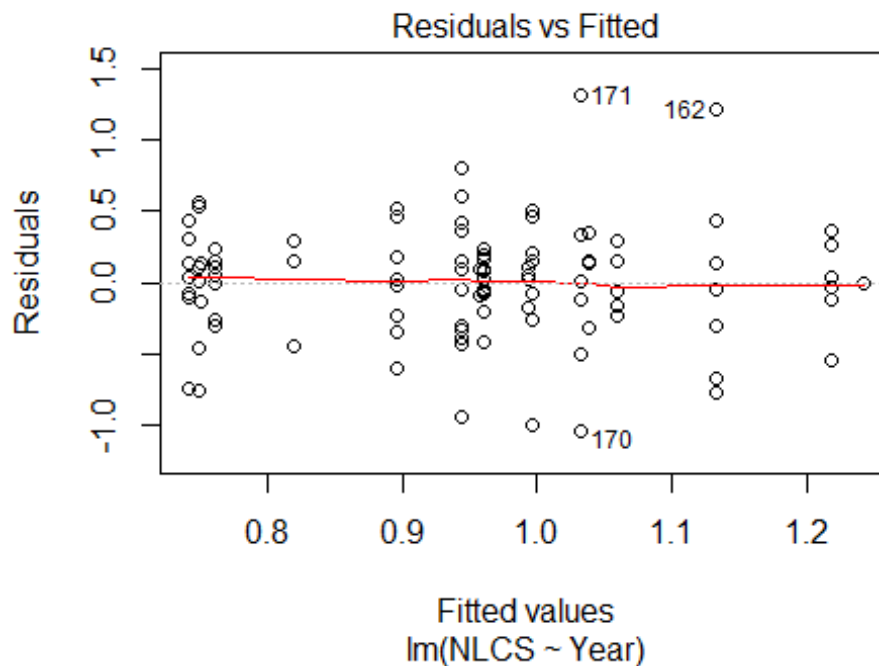
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.33e+00 -2.70e-01 7.22e-16 3.23e-01 1.19e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.1210 0.0741 28.62 < 2e-16 ***
## LastAuthorFemale1 -0.1753 0.1285 -1.36 0.17730
## Year1998 -0.7664 0.1664 -4.61 2.1e-05 ***
## Year1999 -1.1160 0.0741 -15.06 < 2e-16 ***
## Year2000 -0.4900 0.0741 -6.61 1.0e-08 ***
## Year2001 -0.7892 0.3100 -2.55 0.01341 *
## Year2002 -0.6806 0.3852 -1.77 0.08214 .
## Year2003 -0.9935 0.1596 -6.23 4.6e-08 ***
## Year2004 -0.3399 0.6130 -0.55 0.58124
## Year2005 -1.1830 0.2039 -5.80 2.4e-07 ***
## Year2006 -0.5880 0.3469 -1.70 0.09506 .
## Year2007 -1.0057 0.3029 -3.32 0.00151 **
```

```

## Year2008          -0.6740      0.1307    -5.16  2.8e-06 ***
## Year2009          -0.7821      0.2037    -3.84  0.00029 ***
## Year2010          -0.5595      0.2136    -2.62  0.01106 *
## Year2011          -1.0326      0.1418    -7.28  7.0e-10 ***
## Year2012          -0.8661      0.1668    -5.19  2.4e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.507
## Multiple R-squared:  0.221, Adjusted R-squared:  0.0205
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.472  0.922  0.960  0.923  0.986  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      1.27e-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: 79"
## [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
##   13   10   12    5   13    8   14    9    6   12    8   12   12    6   13
## 2011 2012
##    6   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    4    2    3    2   11    7    2    7    7    7    6    6   12
## 2011 2012
##    5    8

```

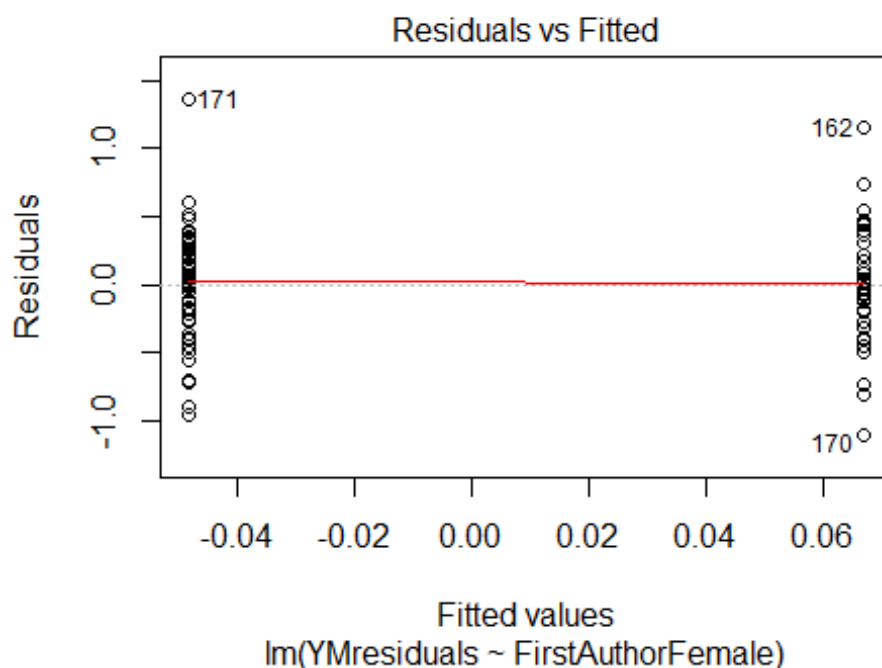
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    3    1    3    2   11    7    2    6    7    7    5    6   12
## 2011 2012
##    5    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity 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 = 0.42, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 5.29150262212918"
## [1] "Male first author team size 2018 geometric mean: 4.71769398031653"
## 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 = 3.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4"
## [1] "Male last author team size 2018 geometric mean: 5.20681125291209"

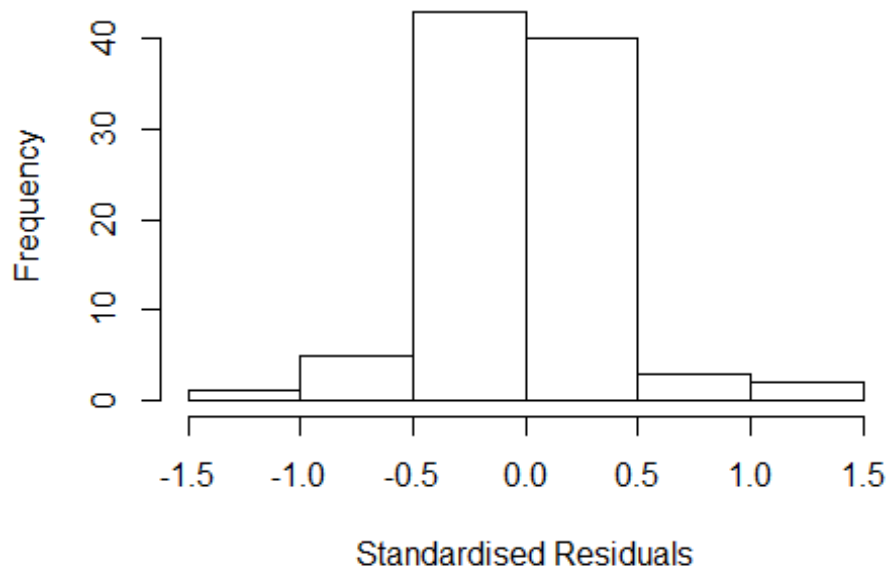
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, 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	-2.545e+15	1	NaN
LastAuthorFemale	4.843e+00	1	2.201
UniqueAuthors	-1.477e+17	4	NaN
Year	-4.277e+17	16	NaN

Residuals from first and last author and team size



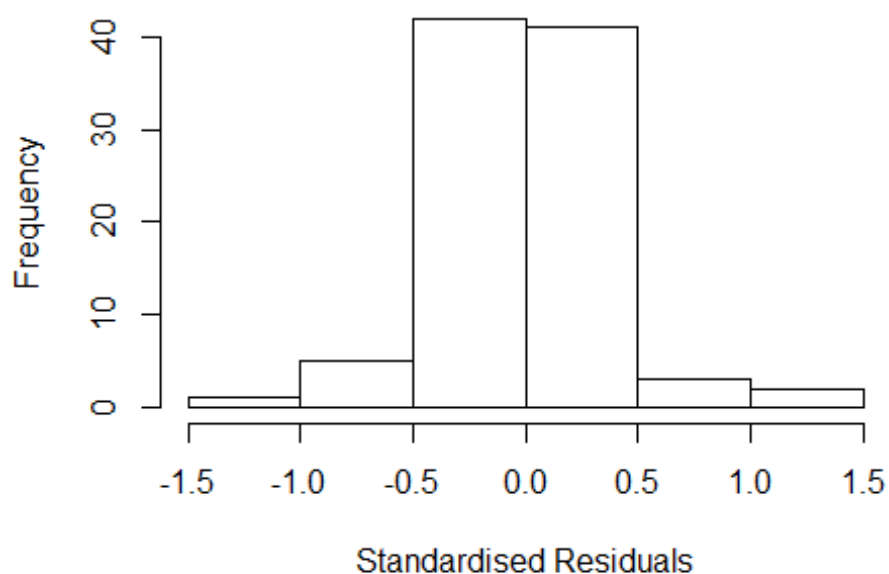
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.10e+00 -2.07e-01 -4.96e-16 2.20e-01 1.14e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.010016 0.146341 6.90 1.8e-09 ***
## FirstAuthorFemale1 0.136548 0.101798 1.34 0.1841
## LastAuthorFemale1 0.149738 0.093912 1.59 0.1153
## UniqueAuthors2 0.238494 0.153449 1.55 0.1246
## UniqueAuthors3 0.052857 0.119214 0.44 0.6588
## UniqueAuthors4 0.144043 0.137960 1.04 0.3000
## UniqueAuthors5 0.095021 0.149808 0.63 0.5279
## Year1997 -0.029397 0.139019 -0.21 0.8331
## Year1998 -0.165760 0.163607 -1.01 0.3144
## Year1999 -0.146016 0.146341 -1.00 0.3218
```

```

## Year2000      -0.277861    0.194453    -1.43    0.1574
## Year2001      -0.374801    0.226968    -1.65    0.1031
## Year2002      -0.303868    0.128028    -2.37    0.0203 *
## Year2003      -0.468544    0.143398    -3.27    0.0017 **
## Year2004        0.000415    0.157319     0.00    0.9979
## Year2005      -0.280576    0.220256    -1.27    0.2069
## Year2006      -0.125674    0.188351    -0.67    0.5068
## Year2007      -0.111165    0.285703    -0.39    0.6984
## Year2008      -0.191862    0.247586    -0.77    0.4410
## Year2009      -0.383526    0.226321    -1.69    0.0945 .
## Year2010      -0.351269    0.188035    -1.87    0.0659 .
## Year2011      -0.221548    0.187301    -1.18    0.2408
## Year2012      -0.278251    0.150942    -1.84    0.0694 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.372
## Multiple R-squared:  0.226, Adjusted R-squared:  -0.014
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 83 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.331  0.909  0.960  0.912  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      1.06e-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 5.749e+14 1      2.398e+07
## LastAuthorFemale  3.156e+00 1      1.776e+00
## Year              1.430e+15 16      2.976e+00

```

Residuals from first and last author



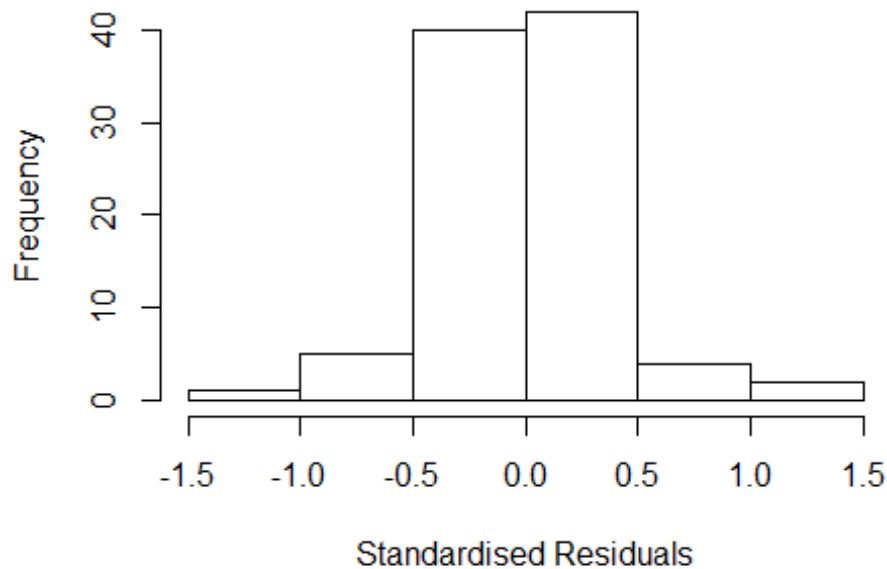
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.01e+00 -1.91e-01  7.63e-16  2.02e-01  1.29e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1010     0.1241   8.87 2.6e-13 ***
## FirstAuthorFemale1  0.1247     0.0982   1.27  0.2082
## LastAuthorFemale1  0.1727     0.0886   1.95  0.0551 .
## Year1997          -0.0433     0.1541  -0.28  0.7796
## Year1998          -0.1140     0.1428  -0.80  0.4272
## Year1999          -0.2370     0.1241  -1.91  0.0600 .
## Year2000          -0.3113     0.2104  -1.48  0.1432
## Year2001          -0.4123     0.1860  -2.22  0.0296 *
## Year2002          -0.2739     0.1287  -2.13  0.0366 *
## Year2003          -0.4648     0.1439  -3.23  0.0018 **
## Year2004           0.0163     0.1295   0.13  0.9000
## Year2005          -0.3498     0.1973  -1.77  0.0803 .
```

```

## Year2006          -0.0934      0.1872   -0.50   0.6193
## Year2007          -0.1702      0.2806   -0.61   0.5461
## Year2008          -0.1074      0.2552   -0.42   0.6750
## Year2009          -0.3871      0.2542   -1.52   0.1320
## Year2010          -0.3047      0.1822   -1.67   0.0987 .
## Year2011          -0.1507      0.1786   -0.84   0.4013
## Year2012          -0.2161      0.1512   -1.43   0.1571
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.375
## Multiple R-squared:  0.187, Adjusted R-squared:  -0.00778
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.211  0.904   0.968   0.913   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      1.06e-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 6.515e+14 1      2.552e+07
## Year              6.515e+14 16      2.904e+00

```


Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.0181 -0.2234 0.0146 0.2368 1.3387
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1415 0.1341 8.51 1.1e-12 ***
## FirstAuthorFemale1 0.1622 0.0905 1.79 0.0771 .
## Year1997 -0.0837 0.1619 -0.52 0.6067
## Year1998 -0.1545 0.1516 -1.02 0.3112
## Year1999 -0.2775 0.1341 -2.07 0.0419 *
## Year2000 -0.3663 0.2070 -1.77 0.0807 .
## Year2001 -0.4716 0.1992 -2.37 0.0204 *
## Year2002 -0.2532 0.1326 -1.91 0.0600 .
## Year2003 -0.4477 0.1441 -3.11 0.0027 **
## Year2004 -0.0617 0.1207 -0.51 0.6105
## Year2005 -0.4011 0.2043 -1.96 0.0533 .
## Year2006 -0.1234 0.1998 -0.62 0.5388
```

```

## Year2007          -0.1799      0.3327   -0.54   0.5904
## Year2008          -0.1372      0.2967   -0.46   0.6451
## Year2009          -0.3927      0.2520   -1.56   0.1232
## Year2010          -0.3131      0.1900   -1.65   0.1035
## Year2011          -0.1657      0.1886   -0.88   0.3823
## Year2012          -0.1971      0.1506   -1.31   0.1944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.389
## Multiple R-squared:  0.157, Adjusted R-squared:  -0.0318
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.213  0.907   0.958   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      1.06e-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 2.636 1          1.624
## Year              2.636 16          1.031
##
## [1] "List of 0 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
## -1.05e+00 -1.86e-01  1.32e-16  2.09e-01  1.40e+00

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1538    0.1226   9.41 2.2e-14 ***
## LastAuthorFemale1 0.2022    0.0836   2.42  0.018 *
## Year1997         -0.0962    0.1527  -0.63  0.531
## Year1998         -0.1668    0.1416  -1.18  0.242
## Year1999         -0.2898    0.1226  -2.36  0.021 *
## Year2000         -0.3193    0.2367  -1.35  0.181
## Year2001         -0.4028    0.1554  -2.59  0.011 *
## Year2002         -0.2844    0.1303  -2.18  0.032 *
## Year2003         -0.4769    0.1393  -3.42  0.001 **
## Year2004          0.0882    0.1226   0.72  0.474
## Year2005         -0.3824    0.1976  -1.93  0.057 .
## Year2006         -0.1009    0.1944  -0.52  0.605
## Year2007         -0.2077    0.2501  -0.83  0.409
## Year2008         -0.1248    0.2462  -0.51  0.614
## Year2009         -0.4293    0.2735  -1.57  0.121
## Year2010         -0.2727    0.1937  -1.41  0.163
## Year2011         -0.1580    0.1683  -0.94  0.351
## Year2012         -0.1977    0.1639  -1.21  0.231
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.178, Adjusted R-squared:  -0.00638
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 82 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.123  0.892  0.955  0.905  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      1.06e-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: 94"
## [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
##    6    3    7    6    1    5    5    9    5    3    8    10    6    10    11
## 2011 2012
##   15   27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    2    6    5    1    3    4    6    4    2    6    3    3    9    11
## 2011 2012
##   14   27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    2    6    4    1    3    3    5    4    2    6    3    3    8    8
## 2011 2012
##   12   26
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.29300356618963"
## [1] "Male first author team size 2018 geometric mean: 4.0356543087298"

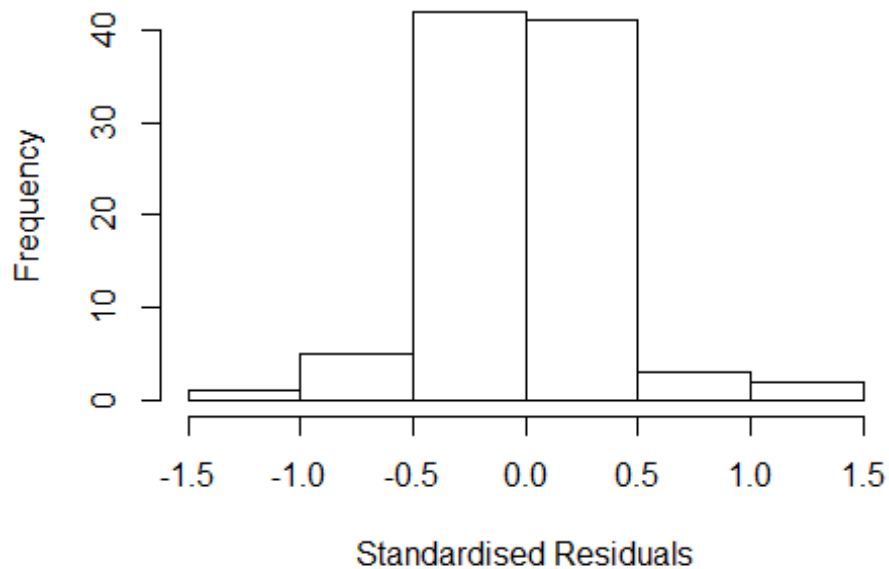
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 30, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.95156811701606"
## [1] "Male last author team size 2018 geometric mean: 3.86911019025"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

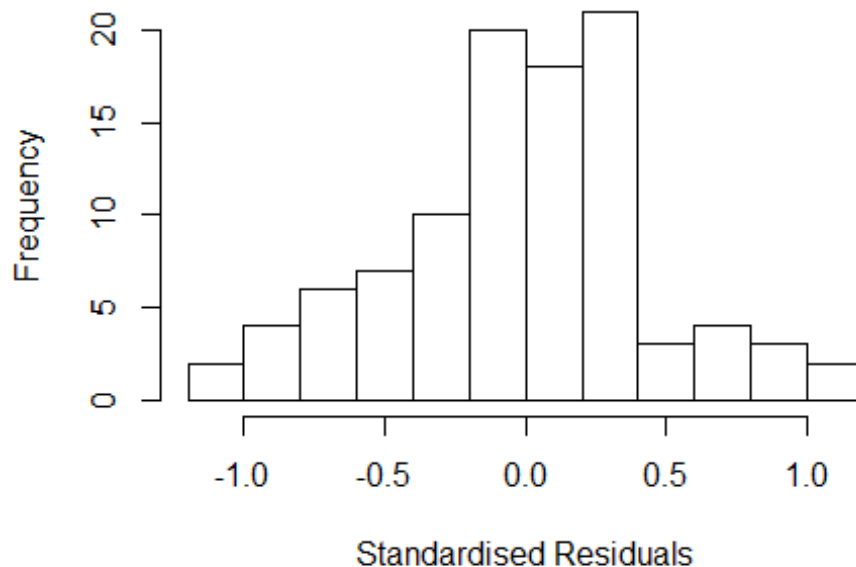
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 34, 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    9.175  1      3.029
## LastAuthorFemale     6.602  1      2.569
## UniqueAuthors    240.625  4      1.985
## Year           1758.336 16      1.263
```

Residuals from first and last author and team size



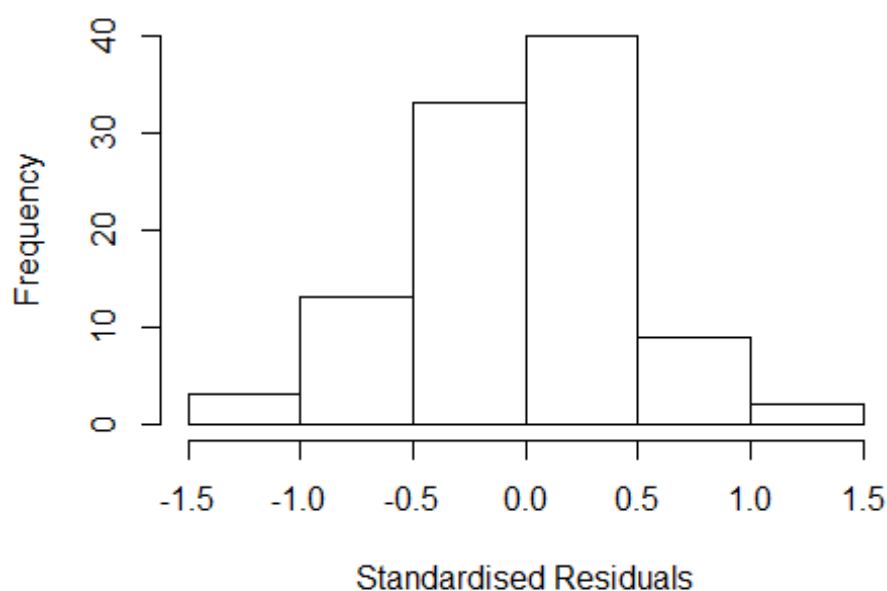
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1281 -0.2584 0.0126 0.2619 1.0872
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.78020 0.14493 5.38 7.7e-07 ***
## FirstAuthorFemale1 -0.00388 0.13324 -0.03 0.97684
## LastAuthorFemale1 0.05444 0.12077 0.45 0.65339
## UniqueAuthors2 0.10502 0.19684 0.53 0.59519
## UniqueAuthors3 0.29992 0.22703 1.32 0.19039
## UniqueAuthors4 0.70225 0.22777 3.08 0.00284 **
## UniqueAuthors5 0.55285 0.14291 3.87 0.00023 ***
## Year1997 0.61058 0.15559 3.92 0.00019 ***
## Year1998 0.07270 0.42677 0.17 0.86518
## Year1999 0.23743 0.19110 1.24 0.21785
```

```

## Year2000      0.96821    0.22753    4.26  5.8e-05 ***
## Year2001     -0.62761    0.44296   -1.42  0.16056
## Year2002      0.34555    0.57441    0.60  0.54922
## Year2003     -0.43909    0.27451   -1.60  0.11380
## Year2004      0.11923    0.26010    0.46  0.64795
## Year2005     -0.16468    0.49106   -0.34  0.73827
## Year2006     -0.02407    0.36882   -0.07  0.94813
## Year2007     -0.03092    0.90762   -0.03  0.97291
## Year2008     -0.01362    0.21286   -0.06  0.94915
## Year2009      0.32061    0.21870    1.47  0.14673
## Year2010     -0.15753    0.24357   -0.65  0.51972
## Year2011     -0.13104    0.14689   -0.89  0.37515
## Year2012     -0.10972    0.17856   -0.61  0.54073
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.462
## Multiple R-squared:  0.349, Adjusted R-squared:  0.163
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 85 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.530  0.854  0.958  0.902  0.987  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      1.00e-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 3.586 1      1.894
## LastAuthorFemale  3.702 1      1.924
## Year              7.342 16      1.064

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.0353 -0.3297 0.0195 0.3163 1.3258
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0123 0.1454 6.96 8.0e-10 ***
## FirstAuthorFemale1 -0.0331 0.1300 -0.25 0.799
## LastAuthorFemale1 0.0263 0.1315 0.20 0.842
## Year1997 0.7361 0.2851 2.58 0.012 *
## Year1998 0.0365 0.2895 0.13 0.900
## Year1999 0.3770 0.1942 1.94 0.056 .
## Year2000 0.8986 0.2049 4.38 3.5e-05 ***
## Year2001 -0.6763 0.3765 -1.80 0.076 .
## Year2002 0.1770 0.4406 0.40 0.689
## Year2003 -0.3166 0.3329 -0.95 0.344
## Year2004 0.2126 0.3552 0.60 0.551
## Year2005 0.0502 1.0039 0.05 0.960
```

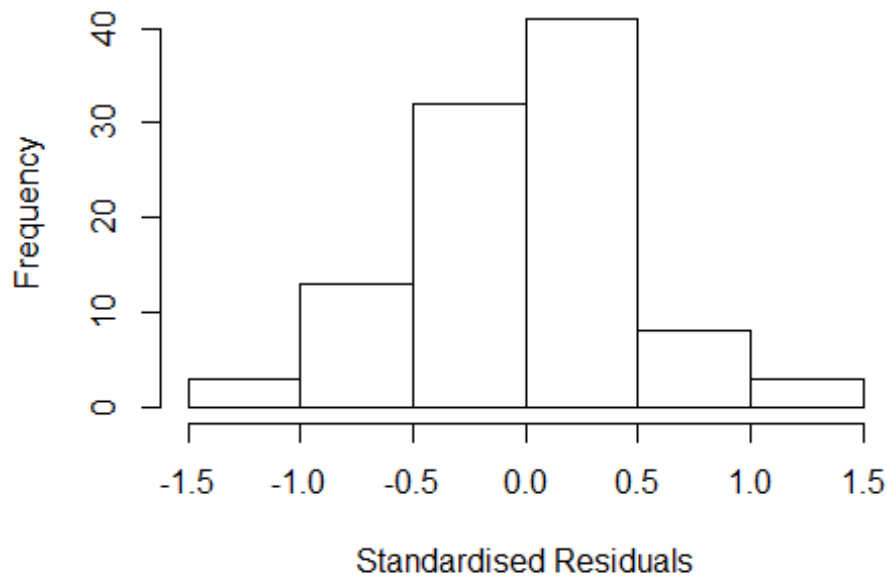


```

## Year2006          -0.0959      0.3926   -0.24    0.808
## Year2007          -0.1157      0.6949   -0.17    0.868
## Year2008           0.1632      0.2937    0.56    0.580
## Year2009           0.3190      0.2349    1.36    0.178
## Year2010          -0.0776      0.2620   -0.30    0.768
## Year2011          -0.0420      0.2105   -0.20    0.842
## Year2012          -0.0110      0.1963   -0.06    0.956
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.526
## Multiple R-squared:  0.182, Adjusted R-squared:  0.00083
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.504  0.854  0.958  0.909  0.981  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      1.00e-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 3.292 1      1.814
## Year              3.292 16      1.038

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.05506 -0.32105  0.00862  0.31352  1.30727
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.01685    0.14162   7.18 2.9e-10 ***
## FirstAuthorFemale1 -0.02746    0.12566  -0.22  0.828
## Year1997        0.74188    0.28803   2.58  0.012 *
## Year1998        0.04419    0.28735   0.15  0.878
## Year1999        0.38455    0.18453   2.08  0.040 *
## Year2000        0.91461    0.18169   5.03 2.8e-06 ***
## Year2001       -0.67145    0.37925  -1.77  0.080 .
## Year2002        0.18416    0.43582   0.42  0.674
## Year2003       -0.30538    0.32422  -0.94  0.349
## Year2004        0.21296    0.35227   0.60  0.547
## Year2005        0.05311    0.97584   0.05  0.957
## Year2006       -0.08315    0.39430  -0.21  0.834
```

```

## Year2007      -0.11908    0.69591   -0.17    0.865
## Year2008      0.17142    0.27812    0.62    0.539
## Year2009      0.33421    0.21276    1.57    0.120
## Year2010     -0.07313    0.26116   -0.28    0.780
## Year2011     -0.03690    0.20286   -0.18    0.856
## Year2012     -0.00265    0.18371   -0.01    0.989
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.183, Adjusted R-squared:  0.0136
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 89 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.508  0.857  0.957  0.909  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      1.00e-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 3.51  1      1.874
## Year              3.51 16      1.040

## [1] "List of 0 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
## -1.0364 -0.3276  0.0286  0.3046  1.3095

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0070    0.1518   6.63 3.2e-09 ***
## LastAuthorFemale1 0.0200    0.1288   0.16  0.877
## Year1997          0.7279    0.2777   2.62  0.010 *
## Year1998          0.0293    0.2932   0.10  0.921
## Year1999          0.3769    0.2007   1.88  0.064 .
## Year2000          0.8769    0.1936   4.53 2.0e-05 ***
## Year2001         -0.7034    0.3694  -1.90  0.060 .
## Year2002          0.1629    0.4073   0.40  0.690
## Year2003         -0.3356    0.3298  -1.02  0.312
## Year2004          0.2018    0.3573   0.56  0.574
## Year2005          0.0254    1.0519   0.02  0.981
## Year2006         -0.0880    0.4022  -0.22  0.827
## Year2007         -0.1456    0.7057  -0.21  0.837
## Year2008          0.1392    0.2838   0.49  0.625
## Year2009          0.3142    0.2360   1.33  0.187
## Year2010         -0.0928    0.2569  -0.36  0.719
## Year2011         -0.0534    0.2111  -0.25  0.801
## Year2012         -0.0225    0.1954  -0.12  0.909
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.508
## Multiple R-squared:  0.185, Adjusted R-squared:  0.0155
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 88 ones are summarized as
##   Min. 1st Qu. Median   Mean 3rd Qu.  Max.
## 0.487 0.852 0.956 0.903 0.981 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      1.00e-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: 100"
## [1] ""

```

```

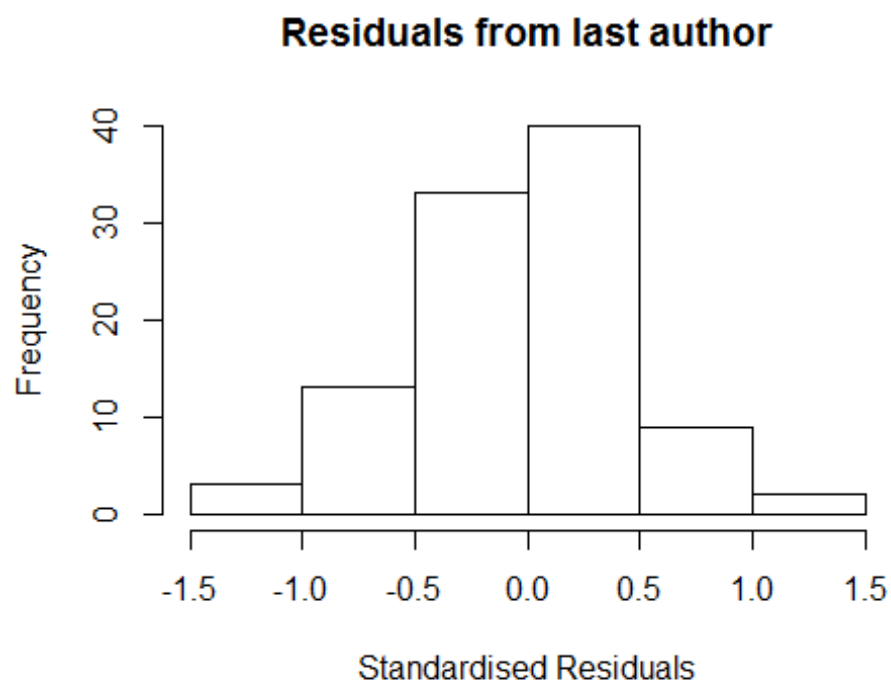
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2718"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    1    2    5    5    2    2    5    7   14    9   10   13   17   23
## 2012
##    17
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    1    0    1    1    2    2    5    8    7    8    9    7   17
## 2012
##    13
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    1    0    1    1    2    1    4    8    7    8    8    6   12
## 2012
##    12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 6.91042323001118"
## [1] "Male first author team size 2018 geometric mean: 3.63424118566428"

## 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 = 9, p-value = 0.08
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 5.0136848842682"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

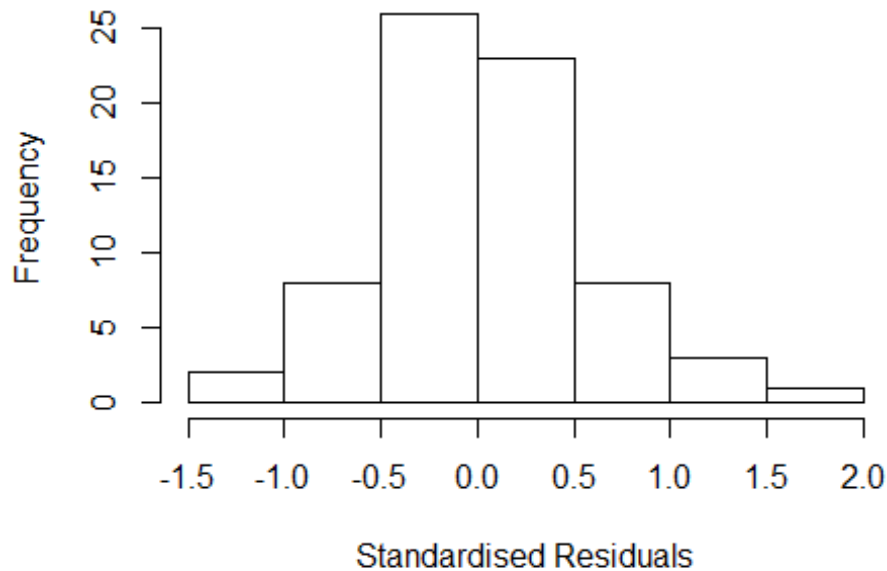
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, 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"
##
```

	GVI	F	Df	$GVI^{1/(2 \cdot Df)}$
FirstAuthorFemale	3.663e+14	1	1	1.914e+07
LastAuthorFemale	3.314e+00	1	1	1.820e+00
UniqueAuthors	2.460e+31	4	4	8.392e+03
Year	8.147e+31	12	12	2.136e+01

Residuals from first and last author and team size



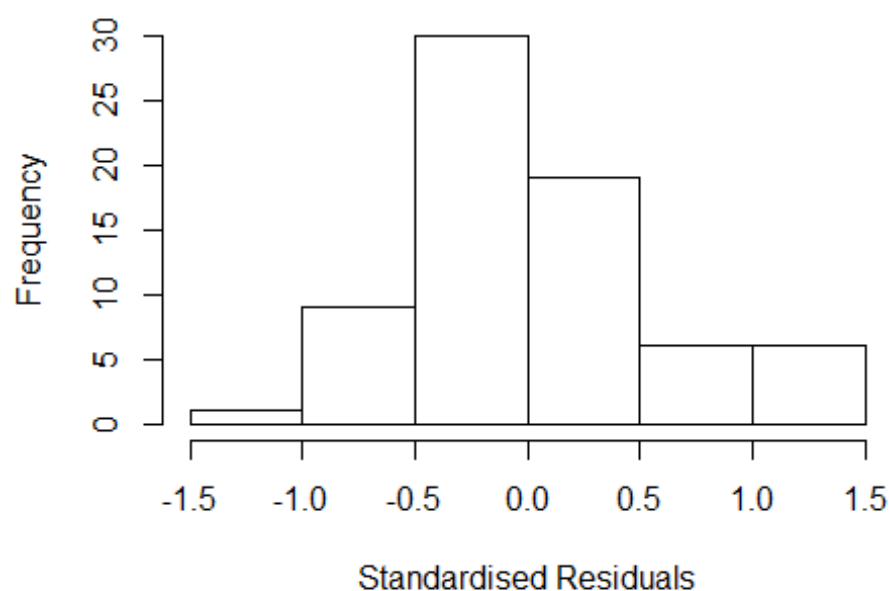
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.04e+00 -3.22e-01 4.44e-16 2.54e-01 1.62e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.86e+00 4.84e-08 3.85e+07 < 2e-16 ***
## FirstAuthorFemale1 -7.23e-02 1.68e-01 -4.30e-01 0.66894
## LastAuthorFemale1 2.71e-01 2.02e-01 1.34e+00 0.18552
## UniqueAuthors2 2.54e-01 3.59e-01 7.10e-01 0.48166
## UniqueAuthors3 3.57e-01 3.49e-01 1.02e+00 0.31081
## UniqueAuthors4 4.46e-01 5.71e-01 7.80e-01 0.43770
## UniqueAuthors5 4.48e-01 3.22e-01 1.39e+00 0.16989
## Year2001 -1.14e+00 3.07e-01 -3.71e+00 0.00051 ***
## Year2002 -2.05e+00 3.02e-01 -6.79e+00 1.1e-08 ***
## Year2003 -1.51e+00 3.99e-01 -3.78e+00 0.00041 ***
```

```

## Year2004      -1.57e+00    5.26e-08 -2.98e+07 < 2e-16 ***
## Year2005      -1.18e+00    9.23e-01 -1.27e+00  0.20849
## Year2006      -9.57e-01    3.17e-01 -3.01e+00  0.00397 **
## Year2007      -1.74e+00    2.66e-01 -6.53e+00  2.7e-08 ***
## Year2008      -1.50e+00    3.94e-01 -3.80e+00  0.00038 ***
## Year2009      -1.55e+00    2.42e-01 -6.39e+00  4.7e-08 ***
## Year2010      -1.67e+00    2.29e-01 -7.27e+00  1.8e-09 ***
## Year2011      -1.35e+00    4.77e-01 -2.83e+00  0.00656 **
## Year2012      -1.40e+00    4.24e-01 -3.31e+00  0.00168 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.342, Adjusted R-squared:  0.114
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~ = 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.207  0.875   0.943   0.881   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           1.41e-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 6.966e+13 1      8.346e+06
## LastAuthorFemale  2.053e+00 1      1.433e+00
## Year              1.404e+14 12      3.886e+00

```


Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.14665 -0.27049 -0.00418 0.22640 1.30677
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.86e+00 3.05e-08 6.10e+07 < 2e-16 ***
## FirstAuthorFemale1 -5.63e-02 1.49e-01 -3.80e-01 0.707
## LastAuthorFemale1 2.41e-01 1.60e-01 1.51e+00 0.137
## Year2001 -7.98e-01 1.49e-01 -5.36e+00 1.6e-06 ***
## Year2002 -1.81e+00 1.49e-01 -1.21e+01 < 2e-16 ***
## Year2003 -1.11e+00 1.02e-01 -1.09e+01 1.7e-15 ***
## Year2004 -1.57e+00 2.86e-08 -5.49e+07 < 2e-16 ***
## Year2005 -7.16e-01 6.77e-01 -1.06e+00 0.295
## Year2006 -6.75e-01 2.92e-01 -2.32e+00 0.024 *
## Year2007 -1.40e+00 1.21e-01 -1.16e+01 < 2e-16 ***
## Year2008 -1.15e+00 2.26e-01 -5.11e+00 4.1e-06 ***
## Year2009 -1.34e+00 3.01e-01 -4.46e+00 4.0e-05 ***
```

```

## Year2010          -1.38e+00   1.93e-01 -7.14e+00   2.0e-09 ***
## Year2011          -9.02e-01   1.52e-01 -5.95e+00   1.8e-07 ***
## Year2012          -1.07e+00   1.71e-01 -6.28e+00   5.3e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.517
## Multiple R-squared:  0.274, Adjusted R-squared:  0.0927
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 63 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.503  0.849  0.944  0.901  0.991  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 2.19e+14 1      1.480e+07
## Year              2.19e+14 12      3.958e+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 + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.07e+00 -2.23e-01 -1.11e-16  2.63e-01  1.44e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)

```

```

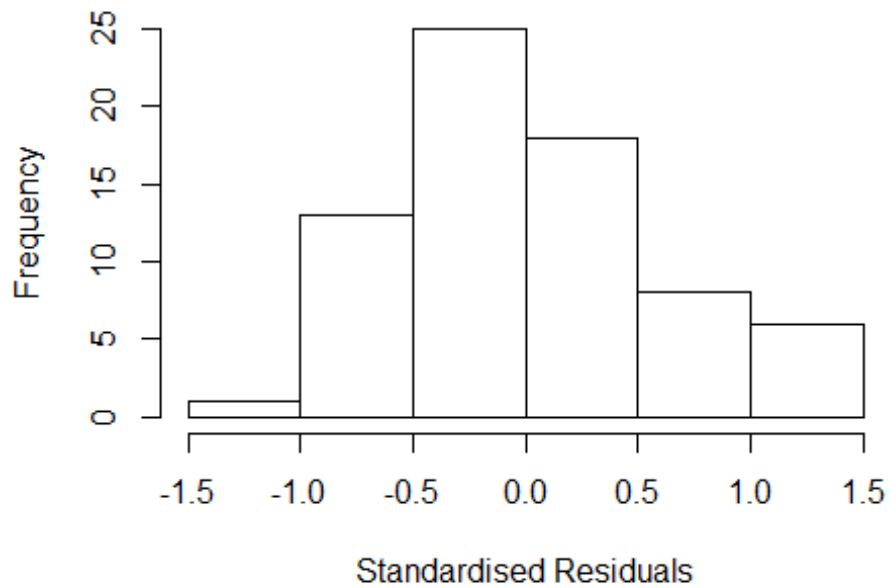
## (Intercept)          1.86e+00    3.21e-08    5.82e+07    < 2e-16 ***
## FirstAuthorFemale1 -1.73e-02    1.48e-01   -1.20e-01     0.91
## Year2001            -8.37e-01    1.48e-01   -5.64e+00    5.6e-07 ***
## Year2002            -1.85e+00    1.48e-01   -1.24e+01    < 2e-16 ***
## Year2003            -1.13e+00    1.12e-01   -1.01e+01    2.5e-14 ***
## Year2004            -1.57e+00    3.58e-08   -4.39e+07    < 2e-16 ***
## Year2005            -6.33e-01    8.15e-01   -7.80e-01     0.44
## Year2006            -6.56e-01    3.12e-01   -2.10e+00     0.04 *
## Year2007            -1.35e+00    1.41e-01   -9.53e+00    2.1e-13 ***
## Year2008            -1.02e+00    1.82e-01   -5.61e+00    6.3e-07 ***
## Year2009            -1.27e+00    2.62e-01   -4.83e+00    1.1e-05 ***
## Year2010            -1.34e+00    1.80e-01   -7.46e+00    5.4e-10 ***
## Year2011            -7.77e-01    1.42e-01   -5.47e+00    1.1e-06 ***
## Year2012            -1.01e+00    1.84e-01   -5.48e+00    1.0e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.507
## Multiple R-squared:  0.247, Adjusted R-squared:  0.0748
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.398  0.820  0.952  0.886  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           1.41e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

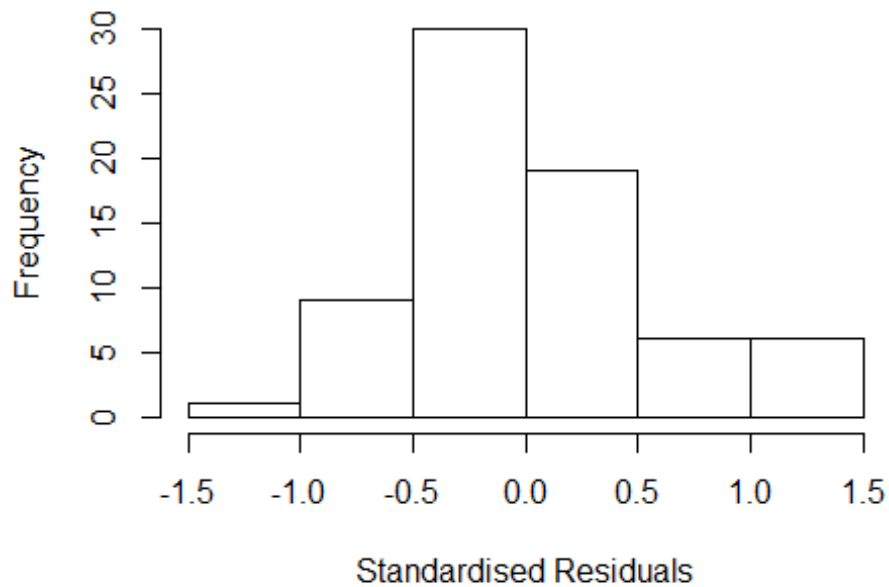
```

Residuals from first author



##	GVIF	Df	$GVIF^{1/(2 \cdot Df)}$
## LastAuthorFemale	NaN	1	NaN
## Year	NaN	12	NaN

Residuals from last author

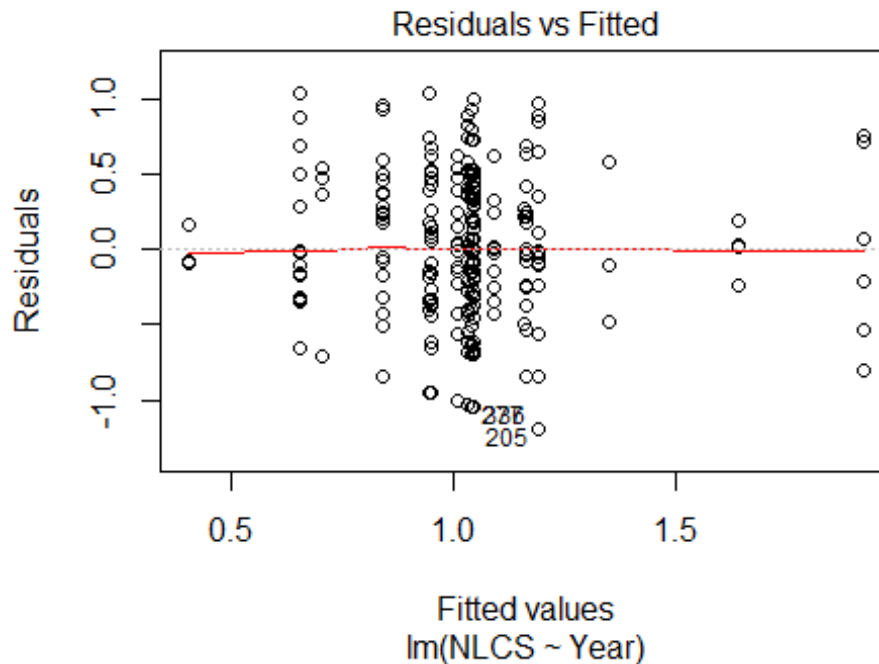


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1738 -0.2609 -0.0182 0.2300 1.2872
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.86e+00 3.05e-08 6.12e+07 < 2e-16 ***
## LastAuthorFemale1 2.31e-01 1.61e-01 1.43e+00 0.158
## Year2001 -8.54e-01 2.20e-08 -3.88e+07 < 2e-16 ***
## Year2002 -1.86e+00 2.86e-08 -6.53e+07 < 2e-16 ***
## Year2003 -1.14e+00 9.01e-02 -1.27e+01 < 2e-16 ***
## Year2004 -1.57e+00 0.00e+00 -Inf < 2e-16 ***
## Year2005 -7.43e-01 6.70e-01 -1.11e+00 0.272
## Year2006 -7.01e-01 2.86e-01 -2.45e+00 0.018 *
## Year2007 -1.41e+00 1.07e-01 -1.32e+01 < 2e-16 ***
## Year2008 -1.19e+00 2.29e-01 -5.19e+00 2.9e-06 ***
## Year2009 -1.38e+00 2.62e-01 -5.27e+00 2.2e-06 ***
## Year2010 -1.39e+00 1.95e-01 -7.11e+00 2.1e-09 ***
## Year2011 -9.21e-01 1.62e-01 -5.69e+00 4.7e-07 ***
## Year2012 -1.10e+00 1.68e-01 -6.53e+00 1.9e-08 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.503
## Multiple R-squared: 0.276, Adjusted R-squared: 0.111
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 64 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.492 0.842 0.942 0.897 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 1.41e-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: 71"
## [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
##    9   19   11    8   16   10   13    6   15   27   17   26   27   27   31
## 2011 2012
##   47   32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6   15   10    3    3    4   10    3    8   19   15   20   21   23   30
## 2011 2012
##   43   27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5   13   10    3    3    4   10    0    8   17   14   17   20   22   29
## 2011 2012
##   39   27
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 14, df = 16, p-value = 0.6

```



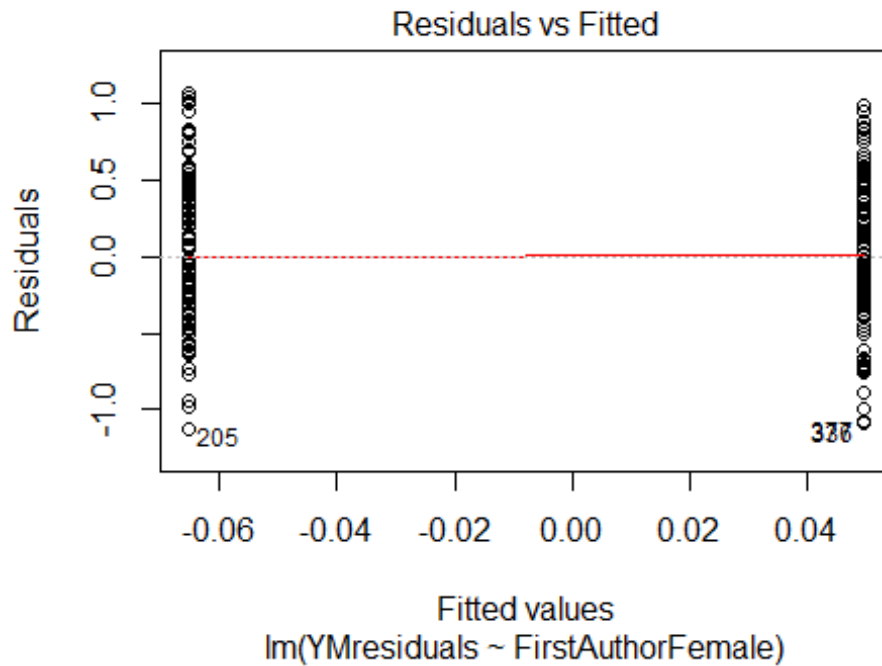
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.023, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 2.94370928043891"
## [1] "Male first author team size 2018 geometric mean: 2.52640240925174"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

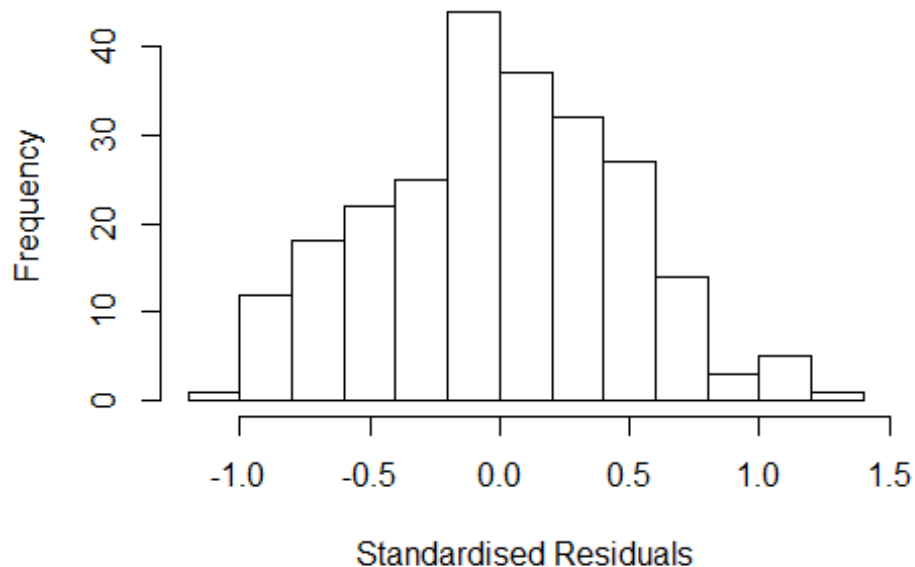
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.1914217624769"
## [1] "Male last author team size 2018 geometric mean: 2.38472745504916"

## 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 = 190, 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 1.745 1          1.321
## LastAuthorFemale  1.488 1          1.220
## UniqueAuthors    3.894 4          1.185
## Year              6.391 15         1.064
```


Residuals from first and last author and team size



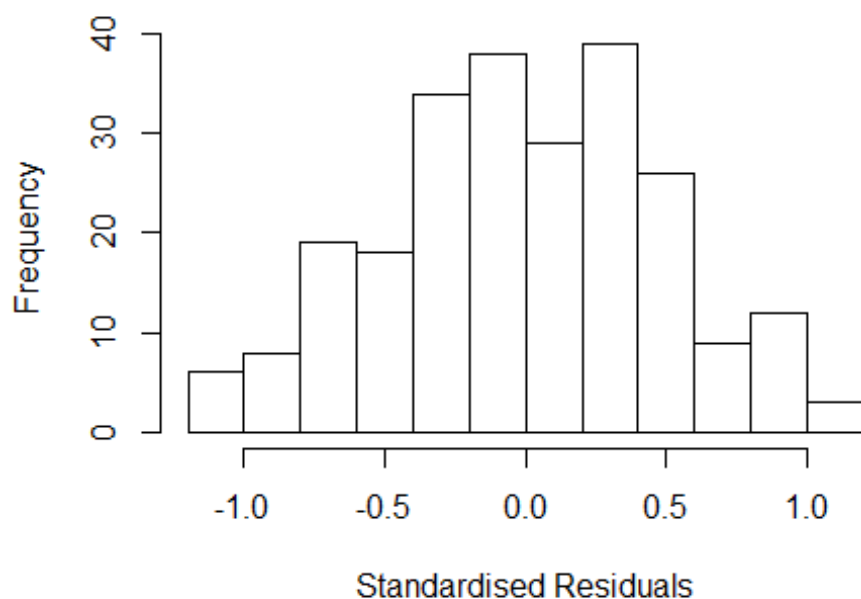
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1642 -0.3531 -0.0135 0.3382 1.2018
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.67511 0.24737 6.77 1.2e-10 ***
## FirstAuthorFemale1 0.10530 0.07655 1.38 0.17034
## LastAuthorFemale1 -0.00693 0.07264 -0.10 0.92410
## UniqueAuthors2 0.26543 0.10009 2.65 0.00859 **
## UniqueAuthors3 0.19064 0.11825 1.61 0.10837
## UniqueAuthors4 0.41600 0.13854 3.00 0.00299 **
## UniqueAuthors5 0.31556 0.09190 3.43 0.00071 ***
## Year1997 -0.96947 0.30027 -3.23 0.00143 **
## Year1998 -1.06714 0.33032 -3.23 0.00143 **
## Year1999 -0.48069 0.30965 -1.55 0.12201
```

```

## Year2000      -1.30672    0.26320   -4.96  1.4e-06 ***
## Year2001      -0.28753    0.28531   -1.01  0.31467
## Year2002      -0.72169    0.27795   -2.60  0.01006 *
## Year2004      -1.11786    0.32641   -3.42  0.00073 ***
## Year2005      -1.26073    0.27965   -4.51  1.1e-05 ***
## Year2006      -0.81444    0.26882   -3.03  0.00274 **
## Year2007      -1.00832    0.27943   -3.61  0.00038 ***
## Year2008      -0.80491    0.30689   -2.62  0.00933 **
## Year2009      -1.04169    0.29006   -3.59  0.00041 ***
## Year2010      -0.86581    0.27439   -3.16  0.00183 **
## Year2011      -0.82177    0.27222   -3.02  0.00284 **
## Year2012      -0.87472    0.27701   -3.16  0.00181 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.502
## Multiple R-squared:  0.228, Adjusted R-squared:  0.154
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 221 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.546  0.877  0.950  0.917  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      4.15e-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 1.576 1 1.256
## LastAuthorFemale 1.332 1 1.154
## Year 1.719 15 1.018

```

Residuals from first and last author



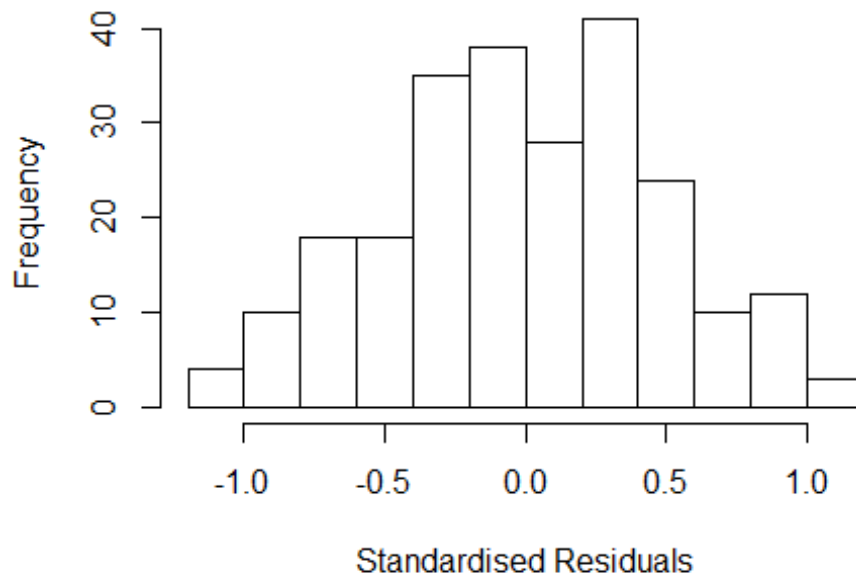
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0999 -0.3289 -0.0148 0.3581 1.0884
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7429 0.2640 6.60 2.9e-10 ***
## FirstAuthorFemale1 0.1162 0.0767 1.52 0.1311
## LastAuthorFemale1 -0.0238 0.0737 -0.32 0.7472
## Year1997 -0.8019 0.2959 -2.71 0.0072 **
## Year1998 -0.8871 0.3313 -2.68 0.0080 **
## Year1999 -0.4657 0.3709 -1.26 0.2107
## Year2000 -1.3726 0.2788 -4.92 1.7e-06 ***
## Year2001 -0.1780 0.2842 -0.63 0.5318
## Year2002 -0.7309 0.2904 -2.52 0.0125 *
## Year2004 -1.0639 0.3602 -2.95 0.0035 **
## Year2005 -1.2088 0.2888 -4.19 4.1e-05 ***
## Year2006 -0.6645 0.2814 -2.36 0.0191 *
```

```

## Year2007          -0.8446      0.2867   -2.95   0.0036 **
## Year2008          -0.6486      0.3106   -2.09   0.0379 *
## Year2009          -0.9092      0.3005   -3.03   0.0028 **
## Year2010          -0.7396      0.2927   -2.53   0.0122 *
## Year2011          -0.7355      0.2833   -2.60   0.0101 *
## Year2012          -0.7536      0.2913   -2.59   0.0103 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.161, Adjusted R-squared:  0.0968
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 218 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.633  0.875   0.954   0.915   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      4.15e-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 1.479 1      1.216
## Year              1.479 15      1.013

```

Residuals from first author



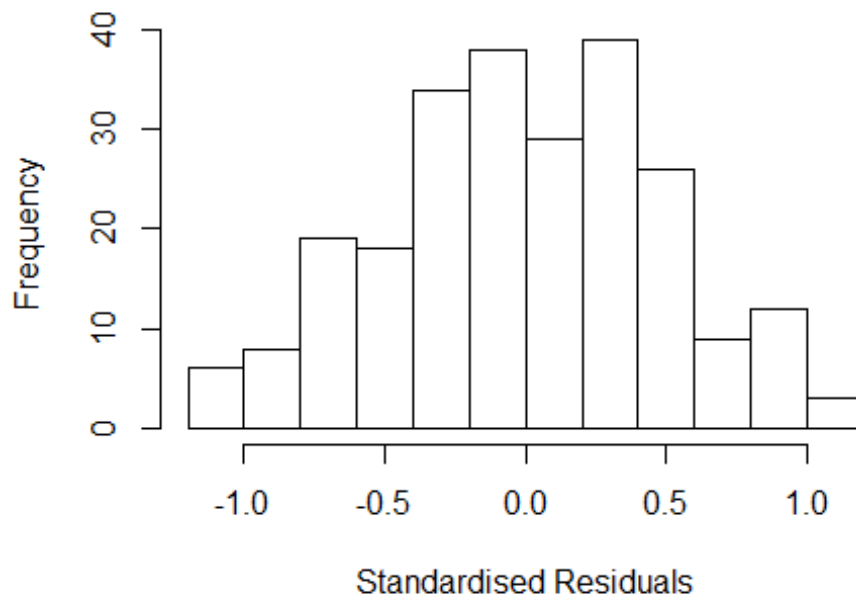
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.105 -0.334 -0.022 0.353 1.072
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7429 0.2641 6.60 2.9e-10 ***
## FirstAuthorFemale1 0.1072 0.0749 1.43 0.1538
## Year1997 -0.8055 0.2951 -2.73 0.0069 **
## Year1998 -0.8973 0.3292 -2.73 0.0069 **
## Year1999 -0.4754 0.3665 -1.30 0.1959
## Year2000 -1.3774 0.2791 -4.93 1.6e-06 ***
## Year2001 -0.1832 0.2815 -0.65 0.5159
## Year2002 -0.7363 0.2903 -2.54 0.0119 *
## Year2004 -1.0695 0.3612 -2.96 0.0034 **
## Year2005 -1.2144 0.2874 -4.23 3.5e-05 ***
## Year2006 -0.6654 0.2816 -2.36 0.0190 *
## Year2007 -0.8501 0.2850 -2.98 0.0032 **
```

```

## Year2008          -0.6560      0.3098   -2.12   0.0353 *
## Year2009          -0.9164      0.2991   -3.06   0.0025 **
## Year2010          -0.7491      0.2889   -2.59   0.0102 *
## Year2011          -0.7451      0.2809   -2.65   0.0086 **
## Year2012          -0.7600      0.2914   -2.61   0.0097 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.161, Adjusted R-squared:  0.101
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 216 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.629  0.873  0.952  0.914  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      4.15e-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 1.284 1          1.133
## Year            1.284 15          1.008

```

Residuals from last author



```
## [1] "List of 0 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
## -1.16271 -0.33383 0.00144 0.35405 1.05497
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7432 0.2639 6.60 2.9e-10 ***
## LastAuthorFemale1 0.0202 0.0731 0.28 0.7821
## Year1997 -0.7609 0.2938 -2.59 0.0102 *
## Year1998 -0.8374 0.3294 -2.54 0.0117 *
## Year1999 -0.4174 0.3828 -1.09 0.2768
## Year2000 -1.3486 0.2744 -4.91 1.7e-06 ***
## Year2001 -0.1138 0.2766 -0.41 0.6811
## Year2002 -0.6701 0.2849 -2.35 0.0195 *
## Year2004 -1.0137 0.3591 -2.82 0.0052 **
## Year2005 -1.1776 0.2891 -4.07 6.4e-05 ***
## Year2006 -0.6281 0.2831 -2.22 0.0275 *
## Year2007 -0.8076 0.2846 -2.84 0.0050 **
```

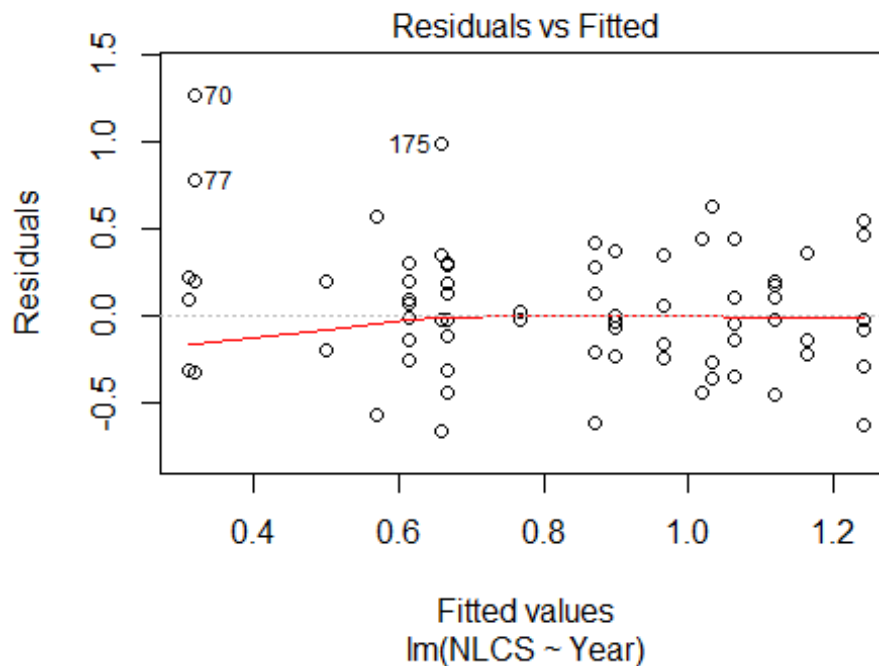
```

## Year2008          -0.5804      0.3037    -1.91    0.0573 .
## Year2009          -0.8767      0.2985    -2.94    0.0037 **
## Year2010          -0.6951      0.2922    -2.38    0.0182 *
## Year2011          -0.6913      0.2814    -2.46    0.0148 *
## Year2012          -0.7187      0.2917    -2.46    0.0145 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.153, Adjusted R-squared:  0.0927
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 214 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.597  0.867  0.948  0.913  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      4.15e-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: 241"
## [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
##    7   15   11    8   25    7   12    3    6    8    8    5    5    7    6
## 2011 2012
##    4    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6    4    2   10    2    8    2    5    8    5    3    3    6    5
## 2011 2012
##    3    2

```



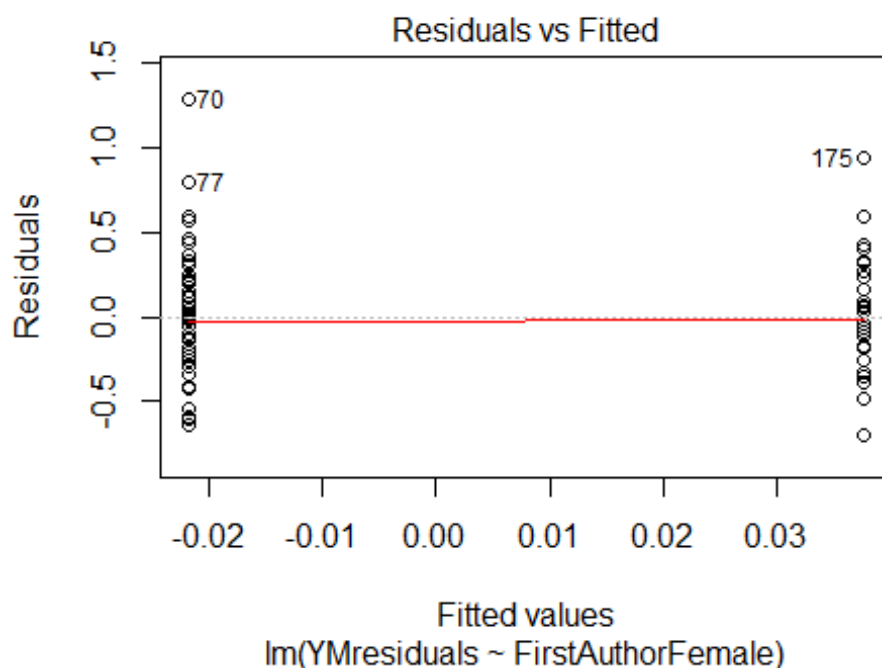
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6    4    2    9    2    8    2    5    8    4    3    2    5    4
## 2011 2012
##    3    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity 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 = 0.012, df = 1, p-value = 0.9
## [1] "Female first author team size 2018 geometric mean: 4.26214759535436"
## [1] "Male first author team size 2018 geometric mean: 4.93242414866094"
## 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 = 4.5, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male last author team size 2018 geometric mean: 5.05478616295142"

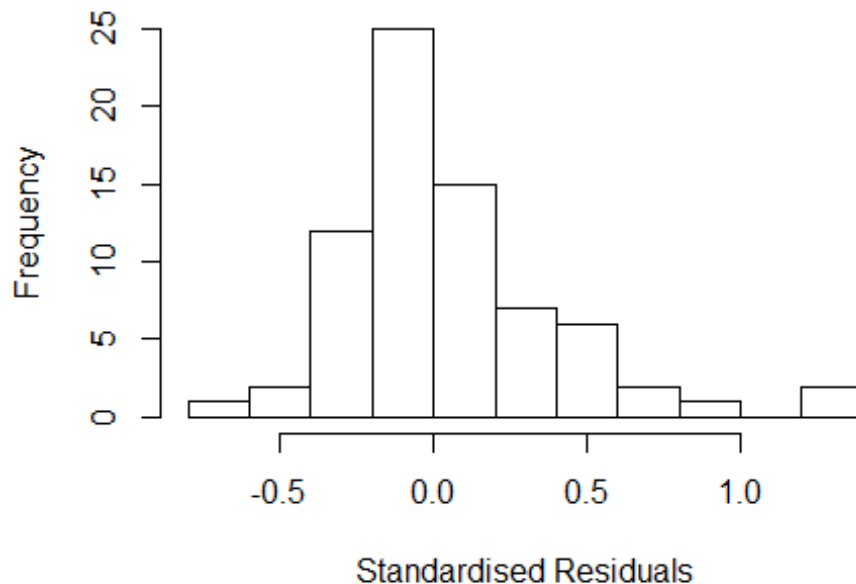
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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] "Regression 1: First author gender, last author gender, team size,
## Year as factors"

##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale    6.979  1      2.642
## LastAuthorFemale   13.495  1      3.674
## UniqueAuthors    7199.825  4      3.035
## Year           41527.963 16      1.394
```

Residuals from first and last author and team size



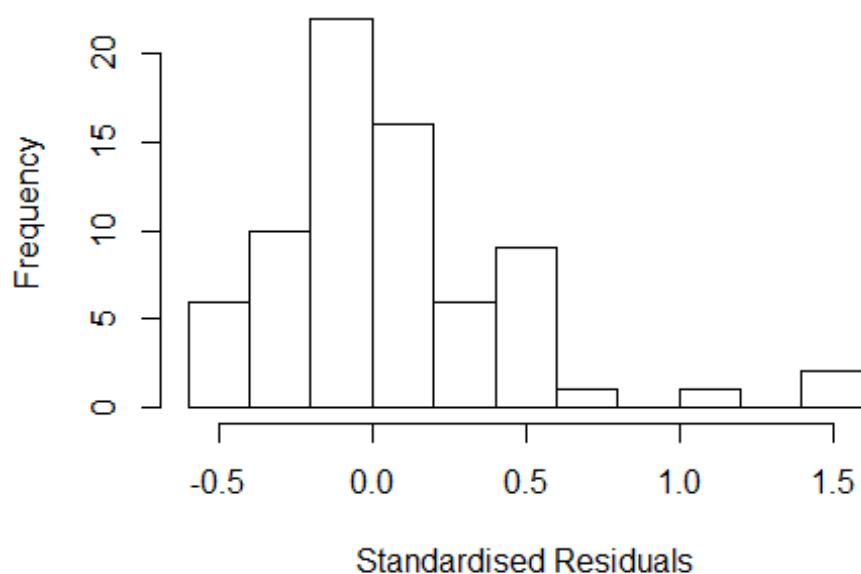
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7111 -0.1742 -0.0198  0.1977  1.3757
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.80865    0.27323     2.96  0.00470 **
## FirstAuthorFemale1 0.10705    0.09936     1.08  0.28645
## LastAuthorFemale1 -0.04894    0.14455    -0.34  0.73635
## UniqueAuthors2   -0.00547    0.21942    -0.02  0.98023
## UniqueAuthors3    0.18849    0.22723     0.83  0.41075
## UniqueAuthors4    0.09226    0.30929     0.30  0.76672
## UniqueAuthors5    0.19822    0.25600     0.77  0.44241
## Year1997          0.33296    0.35807     0.93  0.35691
## Year1998          0.13491    0.18437     0.73  0.46773
## Year1999         -0.11156    0.19616    -0.57  0.57210
```

```

## Year2000      -0.78884    0.29277   -2.69  0.00958 **
## Year2001      0.06411    0.41169    0.16  0.87687
## Year2002     -0.25225    0.19633   -1.28  0.20477
## Year2003     -0.35455    0.76338   -0.46  0.64434
## Year2004      0.15358    0.14569    1.05  0.29687
## Year2005     -0.39361    0.13384   -2.94  0.00495 **
## Year2006     -0.10469    0.29914   -0.35  0.72783
## Year2007      0.08229    0.37128    0.22  0.82550
## Year2008     -0.53750    0.13445   -4.00  0.00021 ***
## Year2009      0.00730    0.20981    0.03  0.97238
## Year2010     -0.64273    0.38733   -1.66  0.10329
## Year2011      0.09301    0.29839    0.31  0.75656
## Year2012     -0.30587    0.15342   -1.99  0.05166 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.31
## Multiple R-squared:  0.572, Adjusted R-squared:  0.383
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 65 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0101 0.8570 0.9590 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          1.37e-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 3.678 1 1.918
## LastAuthorFemale 7.658 1 2.767
## Year 23.679 16 1.104

```

Residuals from first and last author



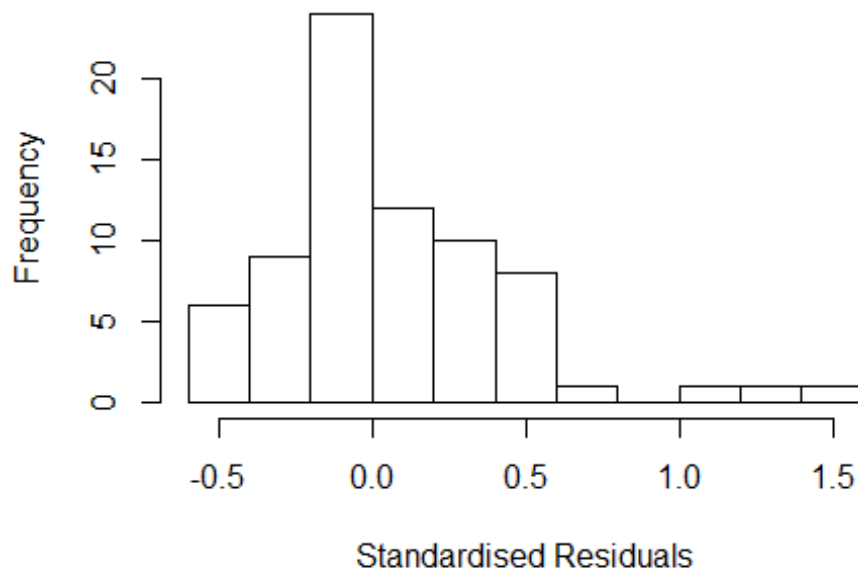
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.5960 -0.1865 -0.0413 0.2178 1.4940
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9812 0.1307 7.51 6.2e-10 ***
## FirstAuthorFemale1 0.0964 0.0950 1.01 0.31500
## LastAuthorFemale1 -0.0632 0.1226 -0.52 0.60844
## Year1997 0.2252 0.2394 0.94 0.35111
## Year1998 0.0104 0.1718 0.06 0.95193
## Year1999 -0.1806 0.1372 -1.32 0.19350
## Year2000 -0.8912 0.1743 -5.11 4.3e-06 ***
## Year2001 0.0368 0.4960 0.07 0.94110
## Year2002 -0.3224 0.1503 -2.14 0.03651 *
## Year2003 -0.3796 0.8918 -0.43 0.67205
## Year2004 0.1129 0.1806 0.63 0.53456
## Year2005 -0.4196 0.1360 -3.09 0.00320 **
```

```

## Year2006          -0.1262      0.2748   -0.46   0.64800
## Year2007          -0.0337      0.3802   -0.09   0.92965
## Year2008          -0.5167      0.1393   -3.71   0.00049 ***
## Year2009          -0.0758      0.1760   -0.43   0.66831
## Year2010          -0.7885      0.3175   -2.48   0.01613 *
## Year2011           0.0908      0.2180    0.42   0.67854
## Year2012          -0.2802      0.1307   -2.14   0.03661 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.326
## Multiple R-squared:  0.551, Adjusted R-squared:  0.401
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0019 0.8560 0.9650 0.8920 0.9900 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 3.611 1      1.900
## Year              3.611 16      1.041

```

Residuals from first author



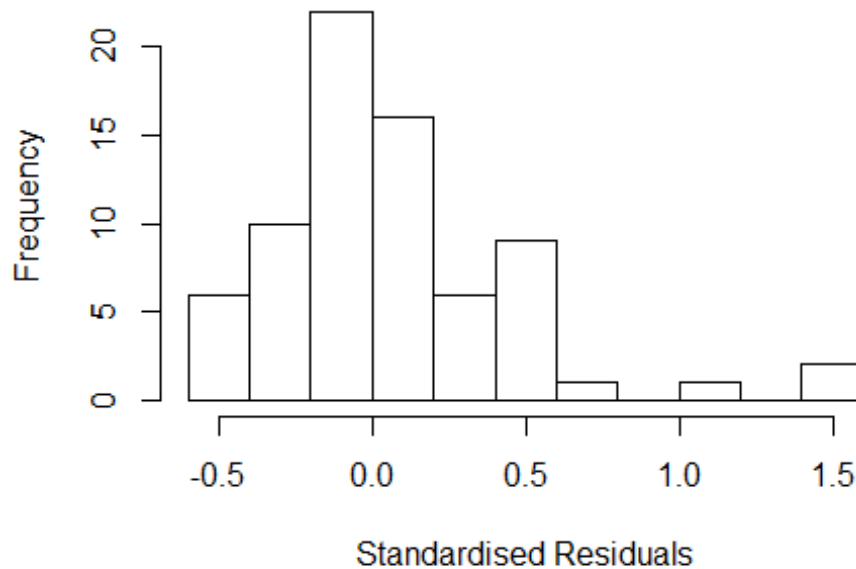
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.589 -0.174 -0.018 0.230 1.521
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9842 0.1312 7.50 5.7e-10 ***
## FirstAuthorFemale1 0.0919 0.0950 0.97 0.33727
## Year1997 0.2234 0.2436 0.92 0.36296
## Year1998 -0.0256 0.1809 -0.14 0.88815
## Year1999 -0.2152 0.1318 -1.63 0.10838
## Year2000 -0.9213 0.1659 -5.55 8.4e-07 ***
## Year2001 0.0338 0.5174 0.07 0.94808
## Year2002 -0.3329 0.1507 -2.21 0.03131 *
## Year2003 -0.4142 1.3710 -0.30 0.76373
## Year2004 0.0998 0.1738 0.57 0.56803
## Year2005 -0.4205 0.1365 -3.08 0.00323 **
## Year2006 -0.1488 0.2775 -0.54 0.59409
```

```

## Year2007          -0.0452      0.3544   -0.13   0.89897
## Year2008          -0.5197      0.1398   -3.72   0.00047 ***
## Year2009          -0.1284      0.1479   -0.87   0.38907
## Year2010          -0.8202      0.3291   -2.49   0.01573 *
## Year2011           0.0753      0.2038    0.37   0.71336
## Year2012          -0.2832      0.1312   -2.16   0.03531 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.314
## Multiple R-squared:  0.558, Adjusted R-squared:  0.421
## Convergence in 38 IRWLS iterations
##
## Robustness weights:
## observation 24 is an outlier with |weight| = 0 ( < 0.0014);
## 8 weights are ~ = 1. The remaining 64 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.014  0.849   0.961   0.894   0.985   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      1.37e-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 10.02 1      3.166
## Year             10.02 16      1.075

```


Residuals from last author



```
## [1] "List of 0 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
## -0.6313 -0.1548 -0.0211 0.2587 1.4842
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0543 0.1255 8.40 2.0e-11 ***
## LastAuthorFemale1 -0.0592 0.1340 -0.44 0.66021
## Year1997 0.1848 0.2462 0.75 0.45616
## Year1998 -0.0652 0.1806 -0.36 0.71952
## Year1999 -0.2557 0.1477 -1.73 0.08900 .
## Year2000 -0.9251 0.1736 -5.33 1.9e-06 ***
## Year2001 -0.0363 0.5185 -0.07 0.94439
## Year2002 -0.3694 0.1658 -2.23 0.03005 *
## Year2003 -0.4547 1.0388 -0.44 0.66331
## Year2004 0.0942 0.1766 0.53 0.59592
## Year2005 -0.4407 0.1443 -3.05 0.00348 **
## Year2006 -0.1641 0.3153 -0.52 0.60493
```

```

## Year2007          -0.1310      0.4141   -0.32   0.75292
## Year2008          -0.5898      0.1344   -4.39   5.2e-05 ***
## Year2009          -0.1172      0.1968   -0.60   0.55384
## Year2010          -0.8402      0.2375   -3.54   0.00083 ***
## Year2011           0.1115      0.2177    0.51   0.61047
## Year2012          -0.3533      0.1255   -2.82   0.00675 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.313
## Multiple R-squared:  0.553, Adjusted R-squared:  0.414
## Convergence in 29 IRWLS iterations
##
## Robustness weights:
## 2 observations c(24,66) are outliers with |weight| <= 0.0002 ( < 0.0014);
## 6 weights are ~1. The remaining 65 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.321  0.861  0.958  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      1.37e-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: 73"
## [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 2007 2008 2009 2010 2011
##    1    1    3    4    5    4    4    3    2    1    5   12    6   10   11
## 2012
##    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011
##    0    0    2    1    0    0    3    2    1    0    3   10    5    7   10

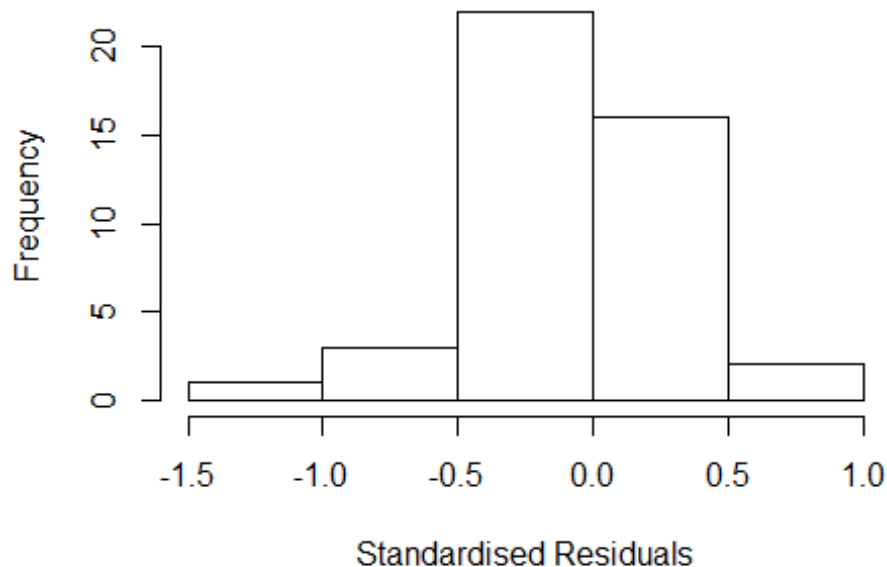
```

```

## 2012
##      4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2007 2008 2009 2010 2011
##      0      0      2      1      0      0      3      2      1      0      3     10      3      7      8
## 2012
##      4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 9"
## [1] "Male first author team size 2018 geometric mean: 4"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 6"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 7.257e+00  1          2.694
## LastAuthorFemale  3.417e+00  1          1.849
## UniqueAuthors    2.741e+16  4         113.432
## Year              4.969e+16 10          6.836

```

Residuals from first and last author and team size



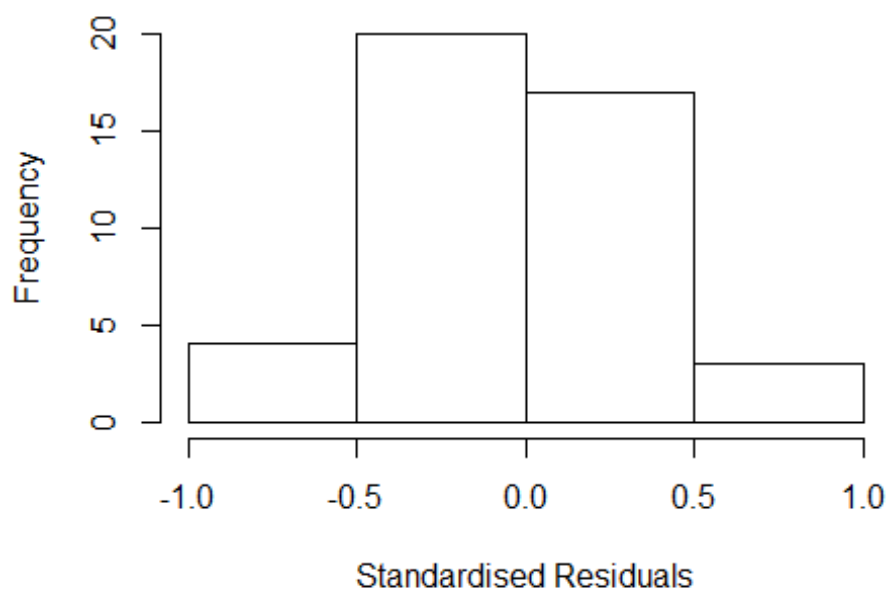
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3783 -0.1655 -0.0324 0.1968 0.6290
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.798 0.339 5.31 1.3e-05 ***
## FirstAuthorFemale1 0.233 0.194 1.20 0.2406
## LastAuthorFemale1 0.014 0.205 0.07 0.9461
## UniqueAuthors2 -0.456 0.321 -1.42 0.1675
## UniqueAuthors3 -0.550 0.299 -1.84 0.0775 .
## UniqueAuthors4 -0.580 0.327 -1.77 0.0877 .
## UniqueAuthors5 -0.353 0.326 -1.08 0.2878
## Year1999 -0.337 0.275 -1.23 0.2297
## Year2002 -0.246 0.219 -1.12 0.2714
## Year2003 -0.173 0.280 -0.62 0.5432
```

```

## Year2004          -0.551      0.154   -3.59   0.0013 **
## Year2007          -0.330      0.268   -1.23   0.2287
## Year2008          -0.306      0.186   -1.65   0.1110
## Year2009          -0.228      0.231   -0.99   0.3314
## Year2010           0.163      0.325    0.50   0.6197
## Year2011          -0.302      0.225   -1.34   0.1899
## Year2012          -0.492      0.395   -1.25   0.2234
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.406, Adjusted R-squared:  0.0545
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.195  0.900  0.977  0.926  0.992  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      2.27e-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 7.505 1      2.740
## LastAuthorFemale  5.676 1      2.382
## Year              20.120 10     1.162

```

Residuals from first and last author



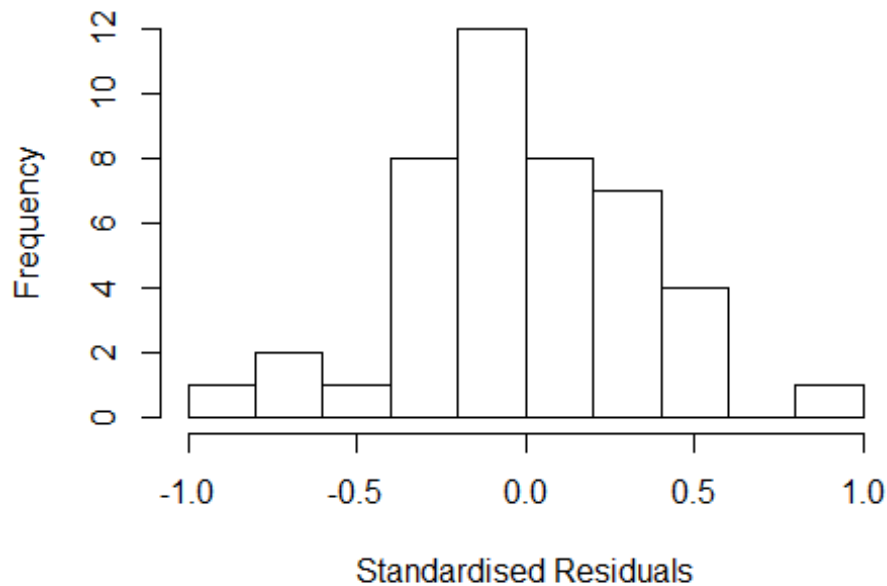
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9848 -0.2297 -0.0104 0.2357 0.9790
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.23869 0.14448 8.57 1.1e-09 ***
## FirstAuthorFemale1 0.18052 0.15696 1.15 0.25890
## LastAuthorFemale1 0.00362 0.20110 0.02 0.98574
## Year1999 -0.23369 0.14448 -1.62 0.11590
## Year2002 0.02204 0.21503 0.10 0.91903
## Year2003 0.23655 0.32251 0.73 0.46879
## Year2004 -0.54169 0.14448 -3.75 0.00073 ***
## Year2007 -0.21173 0.29947 -0.71 0.48483
## Year2008 -0.09673 0.20584 -0.47 0.64169
## Year2009 0.10810 0.22493 0.48 0.63418
## Year2010 0.32911 0.21279 1.55 0.13209
## Year2011 -0.09431 0.17739 -0.53 0.59875
```

```

## Year2012          -0.31889    0.31671   -1.01  0.32179
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.405
## Multiple R-squared:  0.258, Adjusted R-squared:  -0.0288
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.533  0.917  0.960  0.922  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      2.27e-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 3.565 1      1.888
## Year              3.565 10      1.066

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.97944 -0.23355 -0.00772 0.23875 0.97569
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2405 0.1066 11.63 5.0e-13 ***
## FirstAuthorFemale1 0.1812 0.1529 1.19 0.245
## Year1999 -0.2355 0.1066 -2.21 0.035 *
## Year2002 0.0194 0.1925 0.10 0.920
## Year2003 0.2344 0.3083 0.76 0.453
## Year2004 -0.5435 0.1066 -5.10 1.5e-05 ***
## Year2007 -0.2095 0.2756 -0.76 0.453
## Year2008 -0.0952 0.2005 -0.47 0.638
## Year2009 0.1065 0.2046 0.52 0.606
## Year2010 0.3219 0.2203 1.46 0.154
## Year2011 -0.0969 0.1459 -0.66 0.511
## Year2012 -0.3225 0.3142 -1.03 0.312
```

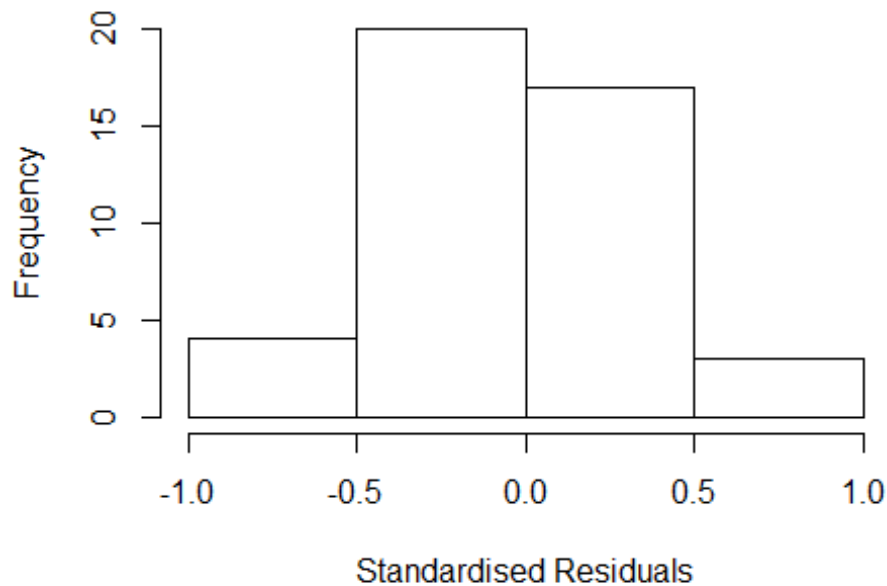


```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.252, Adjusted R-squared:  -0.00548
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.576  0.927  0.963  0.929  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           2.27e-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 3.187 1           1.785
## Year             3.187 10           1.060

```

Residuals from last author



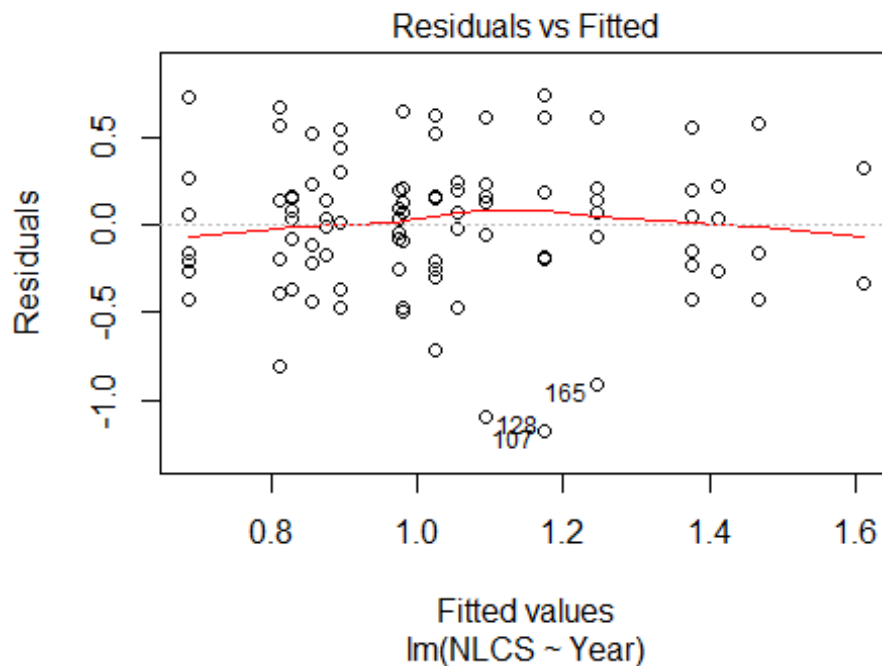
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0156 -0.1945 -0.0215 0.2496 0.9080
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.24083 0.14894 8.33 1.6e-09 ***
## LastAuthorFemale1 -0.00066 0.20293 0.00 0.99742
## Year1999 -0.23583 0.14894 -1.58 0.12315
## Year2002 0.01975 0.21653 0.09 0.92790
## Year2003 0.32467 0.24522 1.32 0.19489
## Year2004 -0.54383 0.14894 -3.65 0.00092 ***
## Year2007 -0.21235 0.30274 -0.70 0.48810
## Year2008 -0.02783 0.19592 -0.14 0.88794
## Year2009 0.16607 0.19923 0.83 0.41072
## Year2010 0.35778 0.18830 1.90 0.06647 .
## Year2011 -0.04959 0.18621 -0.27 0.79170
## Year2012 -0.15643 0.28309 -0.55 0.58440
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.408
## Multiple R-squared:  0.231, Adjusted R-squared:  -0.0333
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.516  0.923  0.966  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           2.27e-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: 44"
## [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
##    11   13    8    8   13    5    7   13   13    8    8   10    6    7    9
## 2011 2012
##    10   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     6    7    5    6    7    2    5    7    6    6    6    6    3    3    7
## 2011 2012
##     7    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     5    7    2    4    7    1    5    6    5    6    4    5    3    3    6
## 2011 2012
##     7    6

```

```
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 25, df = 16, p-value = 0.06
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.8, df = 1, p-value = 0.09

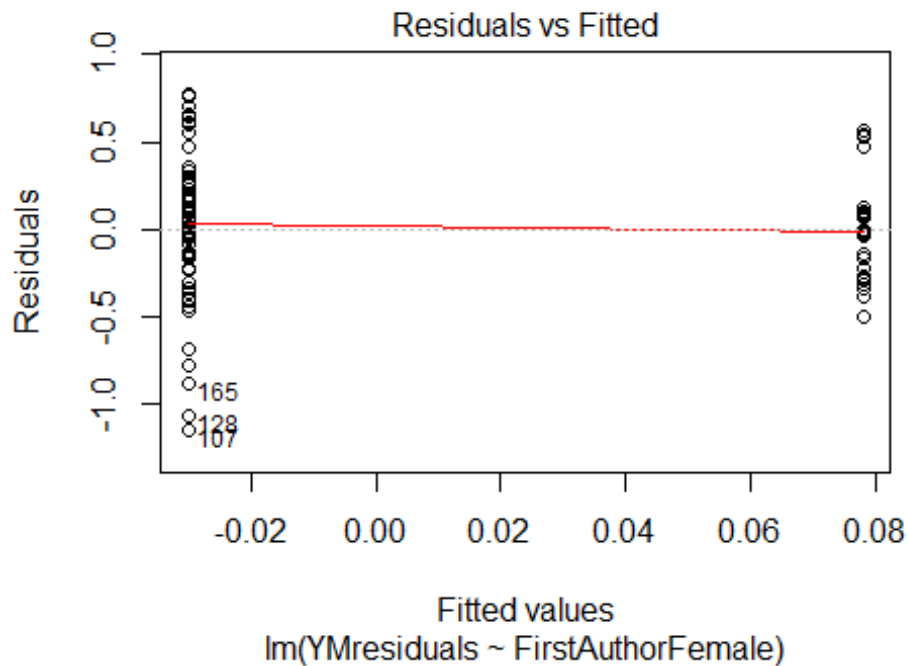
## [1] "Female first author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male first author team size 2018 geometric mean: 2.46621207433047"

## 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 = 2.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
```

```
## [1] "Female last author team size 2018 geometric mean: 3.10723250595386"
## [1] "Male last author team size 2018 geometric mean: 1.73205080756888"

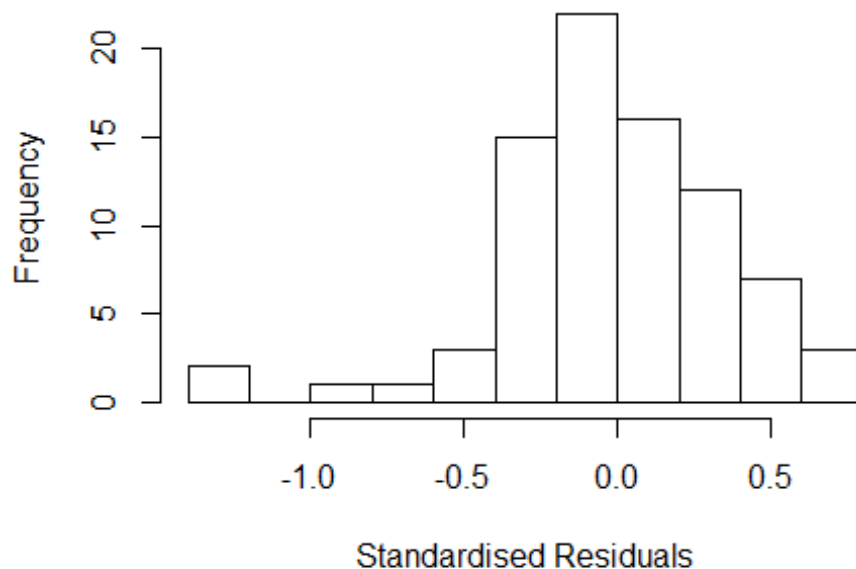
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, 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	1.138e+01	1	3.373
LastAuthorFemale	4.774e+00	1	2.185
UniqueAuthors	8.798e+07	4	9.841
Year	1.251e+09	16	1.924

Residuals from first and last author and team size



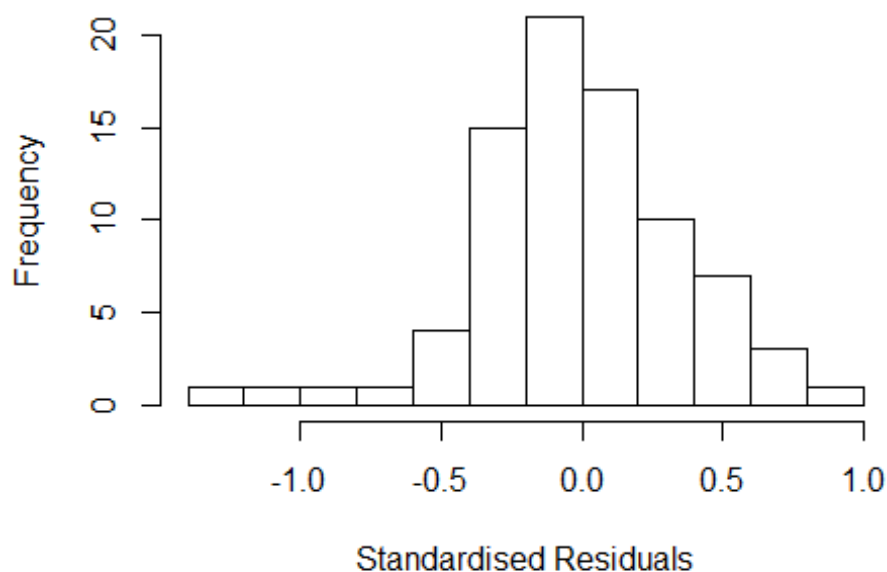
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.33724 -0.21158 -0.00653 0.20708 0.79630
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.79438 0.11987 6.63 1.2e-08 ***
## FirstAuthorFemale1 0.14198 0.15999 0.89 0.378
## LastAuthorFemale1 -0.07169 0.12373 -0.58 0.564
## UniqueAuthors2 0.09713 0.18382 0.53 0.599
## UniqueAuthors3 -0.04860 0.26132 -0.19 0.853
## UniqueAuthors4 -0.17802 0.23821 -0.75 0.458
## UniqueAuthors5 -0.00677 0.34615 -0.02 0.984
## Year1997 0.14542 0.09305 1.56 0.123
## Year1998 0.12381 0.26620 0.47 0.644
## Year1999 0.23267 0.21231 1.10 0.278
```

```

## Year2000      0.11857      0.22598      0.52      0.602
## Year2001      0.53522      0.22614      2.37      0.021 *
## Year2002     -0.01003      0.31753     -0.03      0.975
## Year2003     -0.10515      0.21711     -0.48      0.630
## Year2004     -0.07773      0.33415     -0.23      0.817
## Year2005      0.54964      0.31787      1.73      0.089 .
## Year2006      0.46618      0.28717      1.62      0.110
## Year2007      0.45349      0.44876      1.01      0.316
## Year2008      0.62546      0.32640      1.92      0.060 .
## Year2009      0.53460      0.20978      2.55      0.013 *
## Year2010      0.08819      0.21681      0.41      0.686
## Year2011      0.44282      0.40984      1.08      0.284
## Year2012      0.06809      0.35478      0.19      0.848
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.333
## Multiple R-squared:  0.413, Adjusted R-squared:  0.194
## Convergence in 36 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0704 0.8830 0.9540 0.8850 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      1.22e-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 9.400 1 3.066
## LastAuthorFemale 3.221 1 1.795
## Year 20.627 16 1.099

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.2335 -0.2278 -0.0135  0.2057  0.8485
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8017    0.0846   9.48 9.8e-14 ***
## FirstAuthorFemale1  0.1388    0.1486   0.93  0.354
## LastAuthorFemale1 -0.0388    0.1032  -0.38  0.708
## Year1997         0.1844    0.1032   1.79  0.079 .
## Year1998         0.0888    0.0858   1.03  0.305
## Year1999         0.2192    0.1621   1.35  0.181
## Year2000         0.0748    0.1966   0.38  0.705
## Year2001         0.4793    0.0846   5.67 3.9e-07 ***
## Year2002         0.0131    0.2058   0.06  0.950
## Year2003        -0.2651    0.1344  -1.97  0.053 .
## Year2004        -0.1611    0.2855  -0.56  0.575
## Year2005         0.4318    0.3836   1.13  0.265
```

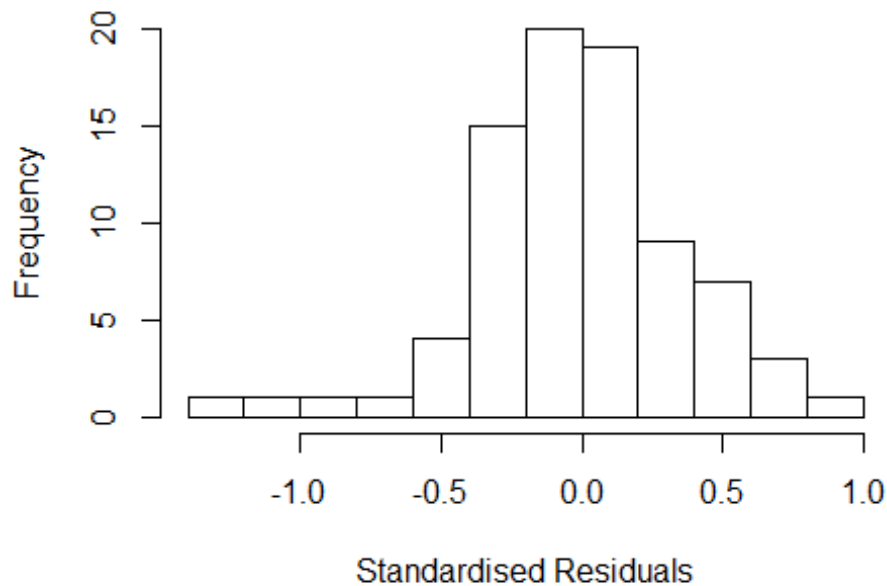


```

## Year2006          0.4338      0.1898      2.29      0.026 *
## Year2007          0.3627      0.2597      1.40      0.167
## Year2008          0.6327      0.3093      2.05      0.045 *
## Year2009          0.5299      0.2092      2.53      0.014 *
## Year2010          0.0165      0.1762      0.09      0.926
## Year2011          0.3931      0.2342      1.68      0.098 .
## Year2012          0.0951      0.2785      0.34      0.734
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.379
## Multiple R-squared:  0.347, Adjusted R-squared:  0.16
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.268  0.884  0.950  0.903  0.988  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      1.22e-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 10.93 1      3.305
## Year              10.93 16      1.078

```

Residuals from first author



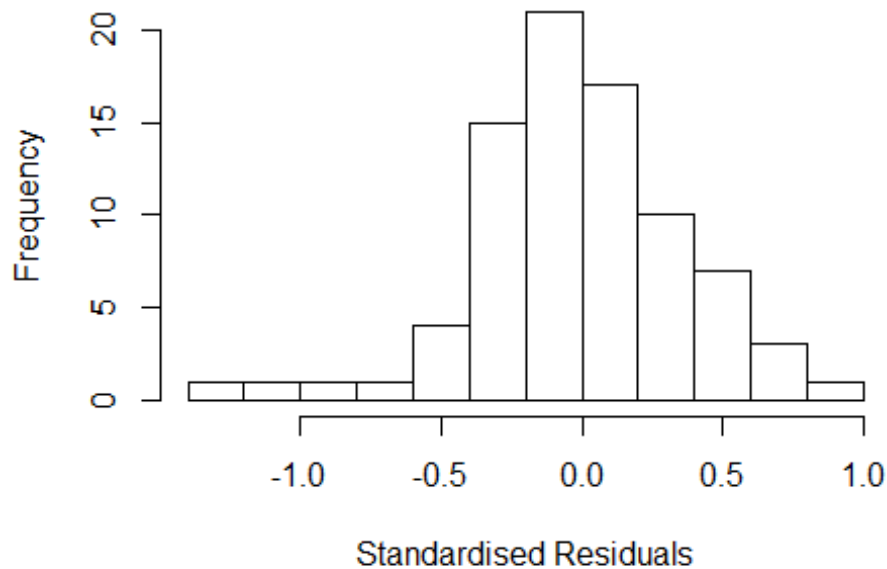
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22217 -0.22892 -0.00401 0.19872 0.86062
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.80182 0.08465 9.47 8.5e-14 ***
## FirstAuthorFemale1 0.13303 0.16019 0.83 0.409
## Year1997 0.17295 0.09819 1.76 0.083 .
## Year1998 0.08868 0.08588 1.03 0.306
## Year1999 0.21945 0.16230 1.35 0.181
## Year2000 0.07545 0.19663 0.38 0.702
## Year2001 0.47918 0.08465 5.66 3.8e-07 ***
## Year2002 0.01355 0.20626 0.07 0.948
## Year2003 -0.26447 0.13460 -1.96 0.054 .
## Year2004 -0.17344 0.27340 -0.63 0.528
## Year2005 0.42035 0.40257 1.04 0.300
## Year2006 0.43799 0.19475 2.25 0.028 *
```

```

## Year2007          0.35723      0.25251      1.41      0.162
## Year2008          0.63084      0.31175      2.02      0.047 *
## Year2009          0.52144      0.21297      2.45      0.017 *
## Year2010          0.00629      0.17076      0.04      0.971
## Year2011          0.38492      0.22326      1.72      0.090 .
## Year2012          0.08865      0.27377      0.32      0.747
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.372
## Multiple R-squared:  0.349, Adjusted R-squared:  0.176
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.258  0.879   0.947   0.901   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      1.22e-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 2.916 1          1.708
## Year              2.916 16          1.034

```

Residuals from last author



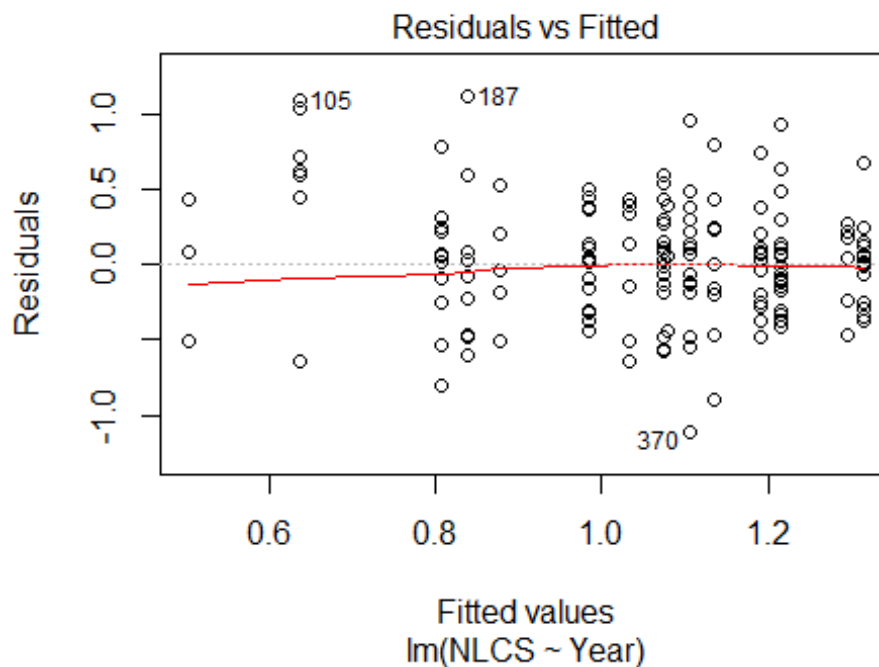
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.360058 -0.230891 -0.000769 0.196179 0.811658
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8019 0.0847 9.47 8.7e-14 ***
## LastAuthorFemale1 -0.0337 0.1135 -0.30 0.7676
## Year1997 0.1827 0.1037 1.76 0.0829 .
## Year1998 0.0886 0.0859 1.03 0.3065
## Year1999 0.2196 0.1625 1.35 0.1811
## Year2000 0.0900 0.2139 0.42 0.6753
## Year2001 0.4791 0.0847 5.66 3.9e-07 ***
## Year2002 0.0431 0.1913 0.23 0.8227
## Year2003 -0.2409 0.1276 -1.89 0.0635 .
## Year2004 -0.1246 0.3045 -0.41 0.6838
## Year2005 0.5581 0.2932 1.90 0.0615 .
## Year2006 0.5405 0.1212 4.46 3.4e-05 ***
```

```

## Year2007          0.4224      0.2207      1.91      0.0601 .
## Year2008          0.6294      0.3137      2.01      0.0490 *
## Year2009          0.6226      0.1397      4.46      3.4e-05 ***
## Year2010          0.0894      0.1553      0.58      0.5666
## Year2011          0.5136      0.1769      2.90      0.0051 **
## Year2012          0.1856      0.2200      0.84      0.4021
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.367
## Multiple R-squared:  0.357, Adjusted R-squared:  0.186
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.140  0.875   0.957   0.898   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      1.22e-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: 82"
## [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
##   19   13   11   17   28   14    9   11    7    7   13   13   17   24   23
## 2011 2012
##   21   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11    6    5   13   13    3    6    9    3    5   10    7   11   18   16
## 2011 2012

```

```
## 15 15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 6 3 12 11 3 2 8 3 4 7 4 8 16 12
## 2011 2012
## 13 13
## [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
```



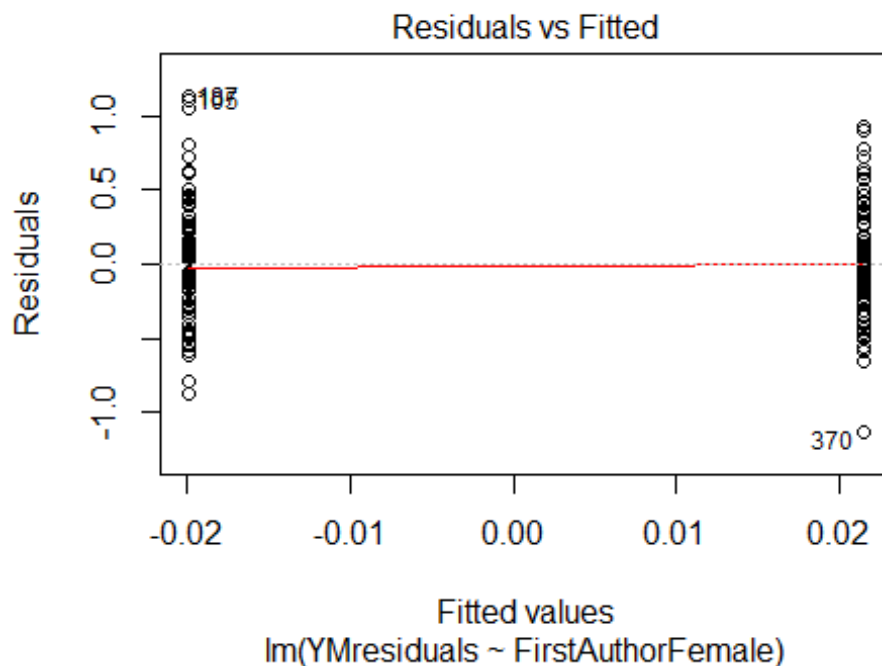
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.09, df = 1, p-value = 0.8

## [1] "Female first author team size 2018 geometric mean: 8.61738083284986"
## [1] "Male first author team size 2018 geometric mean: 4.44690834871737"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

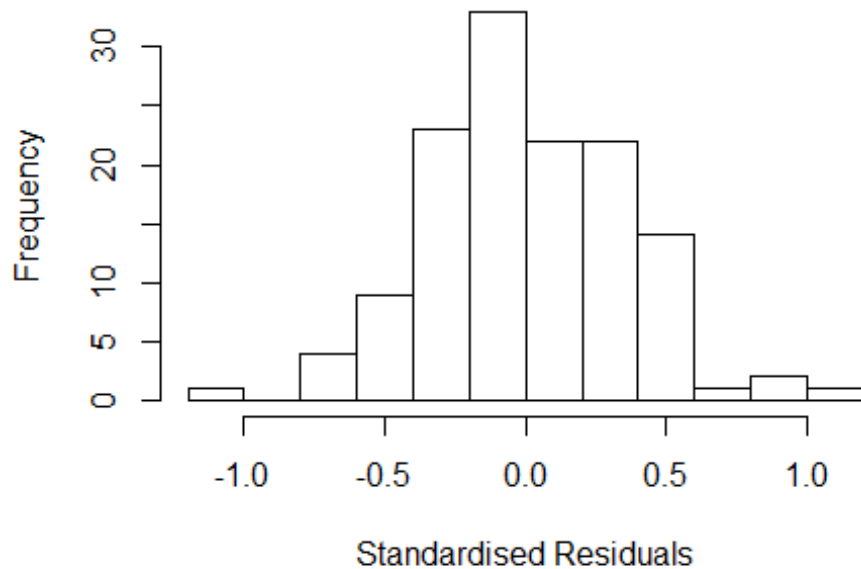
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 7.11378660898013"
## [1] "Male last author team size 2018 geometric mean: 5.20910337958268"

## 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 = 20, 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 3.266 1      1.807
## LastAuthorFemale  1.729 1      1.315
## UniqueAuthors    13.368 4      1.383
## Year              24.797 16     1.106
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.0207 -0.2245 -0.0223 0.2460 1.0567
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8612 0.1909 4.51 1.6e-05 ***
## FirstAuthorFemale1 -0.0364 0.0898 -0.41 0.686
## LastAuthorFemale1 -0.0787 0.0789 -1.00 0.321
## UniqueAuthors2 -0.0213 0.1461 -0.15 0.884
## UniqueAuthors3 0.0198 0.1710 0.12 0.908
## UniqueAuthors4 0.1744 0.1498 1.16 0.247
## UniqueAuthors5 0.3798 0.1757 2.16 0.033 *
## Year1997 0.2692 0.2005 1.34 0.182
## Year1998 0.1238 0.2192 0.56 0.573
## Year1999 0.1591 0.2024 0.79 0.433
```

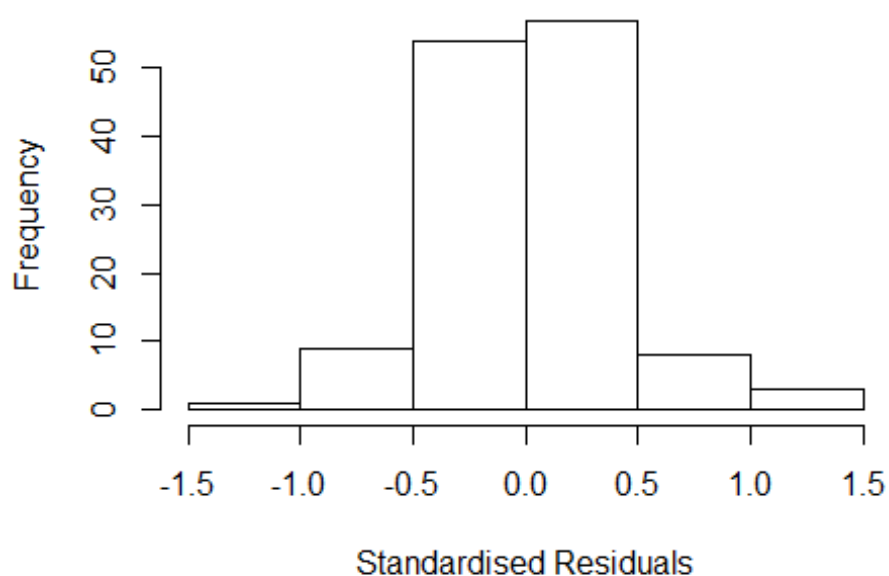


```

## Year2000          -0.4617      0.3631    -1.27      0.206
## Year2001          -0.5432      0.2512    -2.16      0.033 *
## Year2002           0.3449      0.3633      0.95      0.345
## Year2003           0.2188      0.2537      0.86      0.390
## Year2004           0.1004      0.2978      0.34      0.737
## Year2005           0.0334      0.2615      0.13      0.899
## Year2006          -0.3468      0.2415    -1.44      0.154
## Year2007           0.1169      0.2422      0.48      0.630
## Year2008           0.1218      0.2238      0.54      0.587
## Year2009          -0.0368      0.2162    -0.17      0.865
## Year2010          -0.1400      0.2134    -0.66      0.513
## Year2011           0.2472      0.2087      1.18      0.239
## Year2012           0.0215      0.2150      0.10      0.921
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.384, Adjusted R-squared:  0.259
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 121 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.328  0.870  0.949  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      7.58e-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 2.955 1      1.719
## LastAuthorFemale 1.717 1      1.310
## Year              3.697 16      1.042

```

Residuals from first and last author



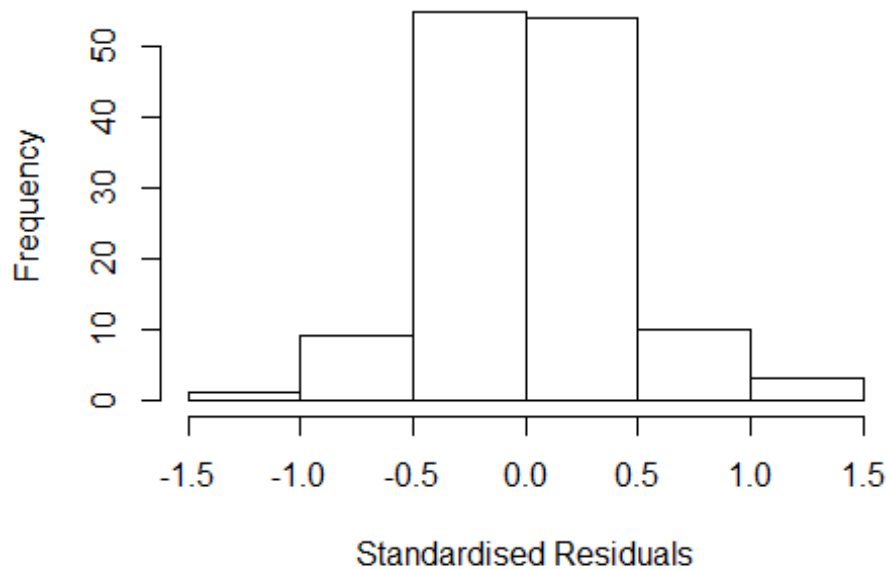
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1962 -0.2546 0.0271 0.2571 1.4005
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9429 0.1540 6.12 1.4e-08 ***
## FirstAuthorFemale1 0.0584 0.0910 0.64 0.522
## LastAuthorFemale1 -0.0710 0.0793 -0.90 0.372
## Year1997 0.2644 0.1724 1.53 0.128
## Year1998 0.0815 0.1715 0.48 0.636
## Year1999 0.1859 0.1713 1.09 0.280
## Year2000 -0.6033 0.6206 -0.97 0.333
## Year2001 -0.4240 0.3056 -1.39 0.168
## Year2002 0.2526 0.3558 0.71 0.479
## Year2003 0.2700 0.2064 1.31 0.193
## Year2004 0.1299 0.2496 0.52 0.604
## Year2005 0.0322 0.2067 0.16 0.877
```

```

## Year2006          -0.3723      0.2260    -1.65      0.102
## Year2007           0.2242      0.2568      0.87      0.384
## Year2008           0.2514      0.2165      1.16      0.248
## Year2009           0.1139      0.1906      0.60      0.551
## Year2010           0.0377      0.1810      0.21      0.835
## Year2011           0.3489      0.1718      2.03      0.045 *
## Year2012           0.1949      0.1761      1.11      0.271
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.34,    Adjusted R-squared:  0.235
## Convergence in 36 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 125 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0897 0.8870 0.9480 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      7.58e-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 2.516 1      1.586
## Year              2.516 16      1.029

```

Residuals from first author



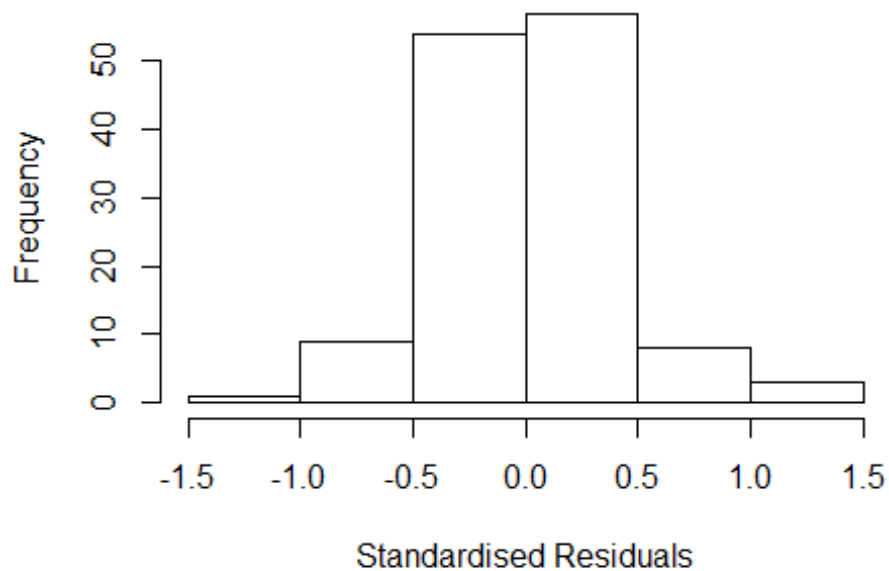
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1631 -0.2530  0.0114  0.2515  1.4041
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9214    0.1490   6.18 1e-08 ***
## FirstAuthorFemale1 0.0488    0.0869    0.56  0.576
## Year1997        0.2682    0.1609    1.67  0.098 .
## Year1998        0.0820    0.1631    0.50  0.616
## Year1999        0.1873    0.1661    1.13  0.262
## Year2000       -0.5934    0.6451   -0.92  0.360
## Year2001       -0.4022    0.3040   -1.32  0.189
## Year2002        0.2741    0.3565    0.77  0.444
## Year2003        0.2949    0.2026    1.46  0.148
## Year2004        0.1389    0.2659    0.52  0.602
## Year2005        0.0414    0.2073    0.20  0.842
## Year2006       -0.3744    0.2246   -1.67  0.098 .
```

```

## Year2007          0.2496      0.2521      0.99      0.324
## Year2008          0.2291      0.2131      1.08      0.285
## Year2009          0.1292      0.1828      0.71      0.481
## Year2010          0.0278      0.1776      0.16      0.876
## Year2011          0.3567      0.1643      2.17      0.032 *
## Year2012          0.1929      0.1716      1.12      0.263
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.334, Adjusted R-squared:  0.234
## Convergence in 37 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 127 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0815 0.8890 0.9490 0.8960 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      7.58e-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 1.479 1      1.216
## Year      1.479 16      1.012

```

Residuals from last author



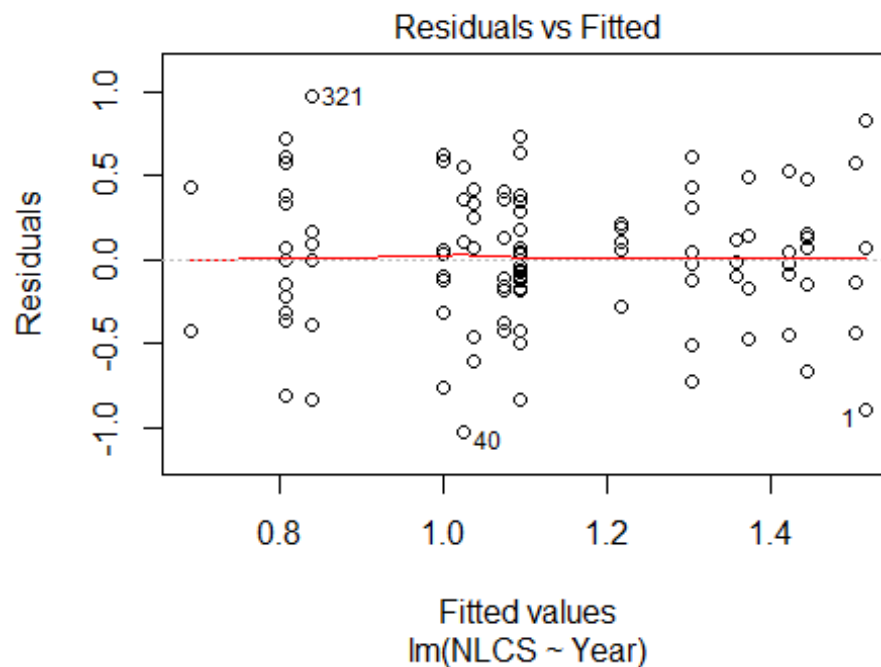
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1578 -0.2601 0.0171 0.2407 1.3498
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9518 0.1484 6.41 3.3e-09 ***
## LastAuthorFemale1 -0.0606 0.0753 -0.81 0.422
## Year1997 0.2803 0.1676 1.67 0.097 .
## Year1998 0.0889 0.1709 0.52 0.604
## Year1999 0.2001 0.1656 1.21 0.230
## Year2000 -0.5720 0.5616 -1.02 0.311
## Year2001 -0.4326 0.3029 -1.43 0.156
## Year2002 0.2437 0.3538 0.69 0.492
## Year2003 0.2821 0.2064 1.37 0.174
## Year2004 0.1524 0.2507 0.61 0.545
## Year2005 0.0650 0.2041 0.32 0.751
## Year2006 -0.3406 0.2126 -1.60 0.112
```

```

## Year2007          0.2359      0.2493      0.95      0.346
## Year2008          0.2853      0.2138      1.33      0.185
## Year2009          0.1420      0.1839      0.77      0.442
## Year2010          0.0496      0.1797      0.28      0.783
## Year2011          0.3621      0.1668      2.17      0.032 *
## Year2012          0.2059      0.1751      1.18      0.242
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.327, Adjusted R-squared:  0.226
## Convergence in 32 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 122 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.116  0.888  0.947  0.892  0.985  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      7.58e-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: 132"
## [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
##   10   25   18   22   13   19   12   19   19   20   16   22   36   20   19
## 2011 2012
##   26   25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    4    2    3    5    8    4    8    8    9    3    6   14    6    6
## 2011 2012

```

```
## 13 6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 4 2 1 3 3 2 5 6 5 3 5 12 5 6
## 2011 2012
## 11 6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 16, df = 16, p-value = 0.5
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.3, df = 1, p-value = 0.2

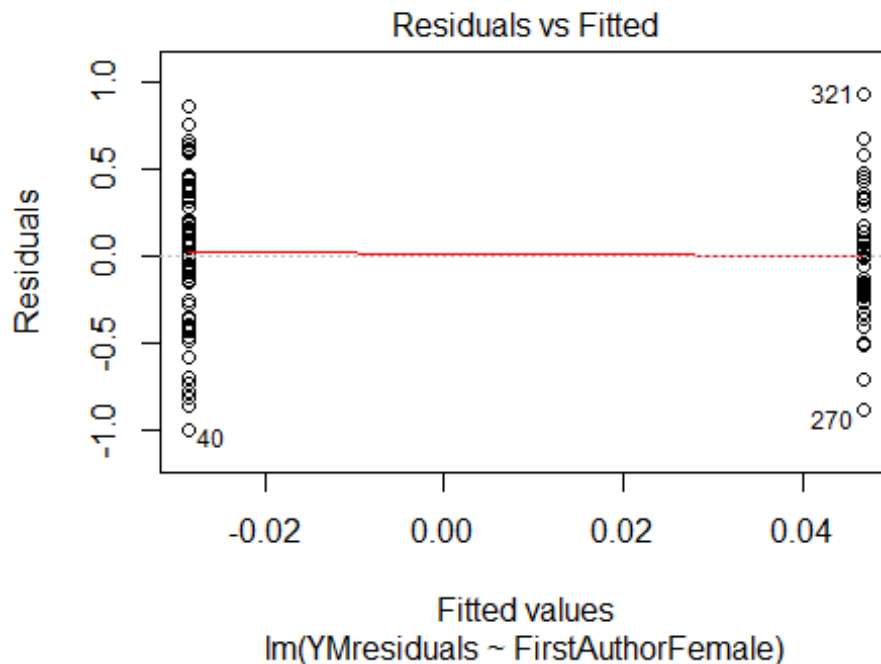
## [1] "Female first author team size 2018 geometric mean: 4.12891791733337"
## [1] "Male first author team size 2018 geometric mean: 3.82592918496298"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



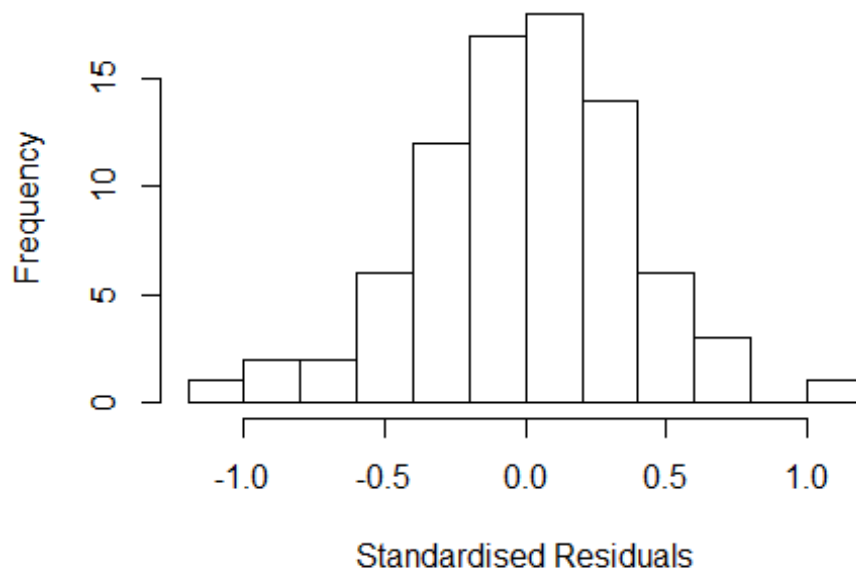
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 26, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.57305051927326"
## [1] "Male last author team size 2018 geometric mean: 3.63539873988721"

## 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 = 35, 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  3.240  1         1.800
## LastAuthorFemale   6.862  1         2.620
## UniqueAuthors    105.842  4         1.791
## Year              702.808 16         1.227
```

Residuals from first and last author and team size



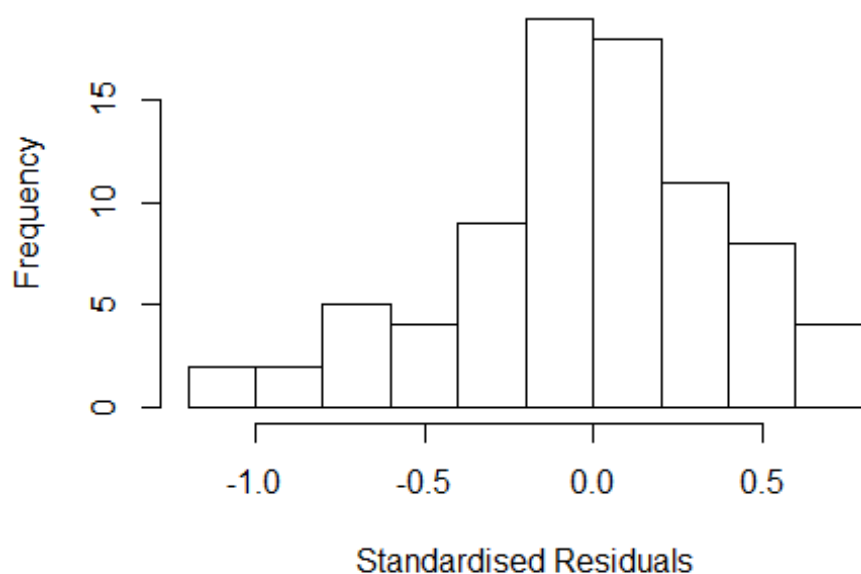
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1005 -0.2392 0.0251 0.2374 1.0227
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3783 0.6382 2.16 0.035 *
## FirstAuthorFemale1 -0.0199 0.0906 -0.22 0.827
## LastAuthorFemale1 0.1742 0.1489 1.17 0.247
## UniqueAuthors2 -0.1894 0.4736 -0.40 0.691
## UniqueAuthors3 -0.0323 0.3660 -0.09 0.930
## UniqueAuthors4 0.1398 0.3358 0.42 0.679
## UniqueAuthors5 0.1702 0.3199 0.53 0.597
## Year1997 -0.2778 0.6840 -0.41 0.686
## Year1998 -0.7577 0.6621 -1.14 0.257
## Year1999 -0.2706 0.5467 -0.49 0.622
```

```

## Year2000          0.0935      0.6383      0.15      0.884
## Year2001         -0.2731      0.7170     -0.38      0.705
## Year2002         -0.4062      0.6764     -0.60      0.550
## Year2003         -0.3498      0.5996     -0.58      0.562
## Year2004         -0.2610      0.6557     -0.40      0.692
## Year2005         -0.2803      0.6231     -0.45      0.655
## Year2006         -0.1286      0.8350     -0.15      0.878
## Year2007         -0.2940      0.6133     -0.48      0.633
## Year2008         -0.4183      0.6003     -0.70      0.489
## Year2009         -0.8321      0.6101     -1.36      0.178
## Year2010         -0.3336      0.5766     -0.58      0.565
## Year2011         -0.7932      0.5985     -1.33      0.190
## Year2012         -0.0749      0.5817     -0.13      0.898
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.363, Adjusted R-squared:  0.125
## Convergence in 30 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.438  0.899  0.959   0.918   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      1.22e-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 2.564 1          1.601
## LastAuthorFemale  4.521 1          2.126
## Year             11.394 16          1.079

```

Residuals from first and last author



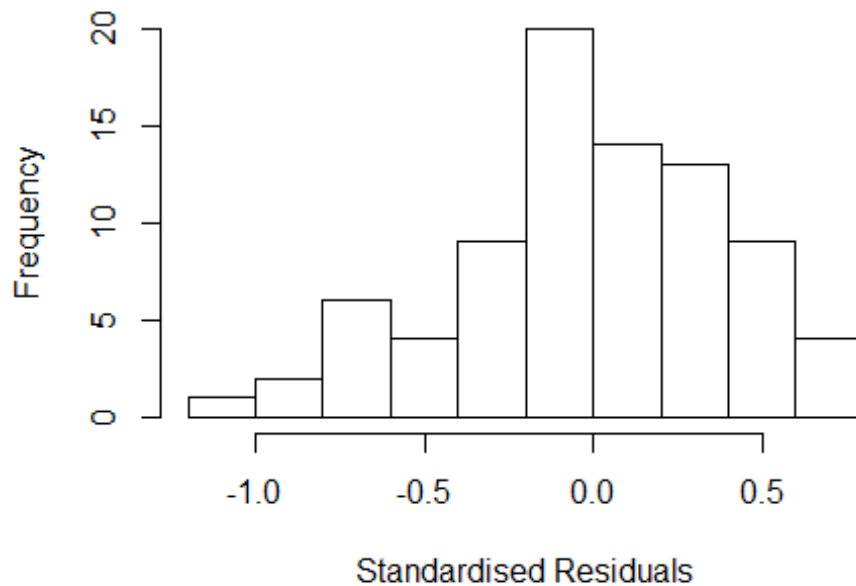
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1340 -0.2119 0.0123 0.2430 0.7297
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5551 0.6840 2.27 0.026 *
## FirstAuthorFemale1 -0.0282 0.0875 -0.32 0.748
## LastAuthorFemale1 0.1558 0.1317 1.18 0.241
## Year1997 -0.4211 0.7589 -0.55 0.581
## Year1998 -0.8656 0.7685 -1.13 0.264
## Year1999 -0.2689 0.6815 -0.39 0.695
## Year2000 -0.0764 0.7330 -0.10 0.917
## Year2001 -0.3232 0.7959 -0.41 0.686
## Year2002 -0.3952 0.7755 -0.51 0.612
## Year2003 -0.3763 0.7216 -0.52 0.604
## Year2004 -0.3051 0.7582 -0.40 0.689
## Year2005 -0.3737 0.7042 -0.53 0.598
```

```

## Year2006          -0.1307      0.7947   -0.16    0.870
## Year2007          -0.4539      0.7301   -0.62    0.536
## Year2008          -0.4658      0.6982   -0.67    0.507
## Year2009          -0.9159      0.7124   -1.29    0.203
## Year2010          -0.3433      0.6950   -0.49    0.623
## Year2011          -0.8648      0.7143   -1.21    0.231
## Year2012          -0.1001      0.6957   -0.14    0.886
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.426
## Multiple R-squared:  0.33,   Adjusted R-squared:  0.138
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.460  0.903  0.967  0.919  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      1.22e-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 2.793 1      1.671
## Year              2.793 16      1.033

```

Residuals from first author



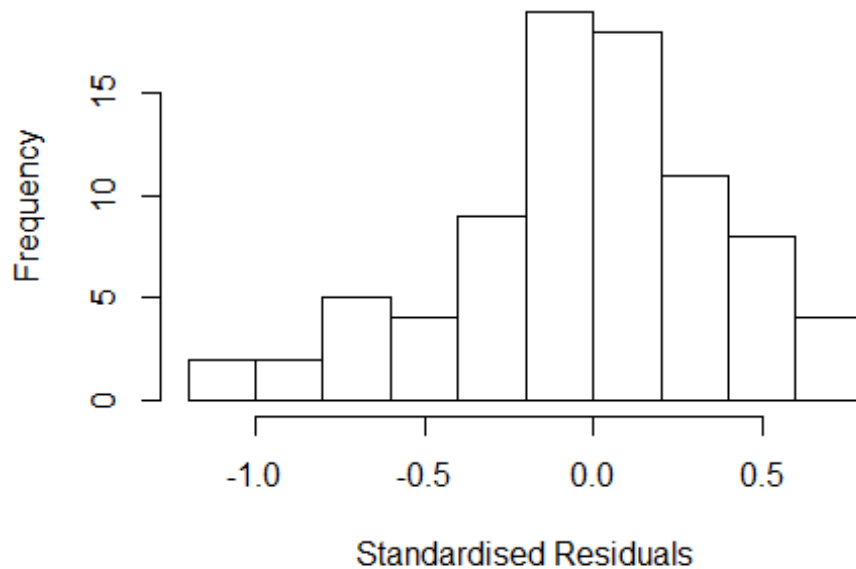
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.176752 -0.215524 -0.000971 0.281916 0.777759
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5622 0.6562 2.38 0.02 *
## FirstAuthorFemale1 -0.0185 0.0889 -0.21 0.84
## Year1997 -0.3855 0.7435 -0.52 0.61
## Year1998 -0.8727 0.7552 -1.16 0.25
## Year1999 -0.2858 0.6627 -0.43 0.67
## Year2000 -0.0890 0.7048 -0.13 0.90
## Year2001 -0.1744 0.6926 -0.25 0.80
## Year2002 -0.3343 0.7040 -0.47 0.64
## Year2003 -0.3248 0.6762 -0.48 0.63
## Year2004 -0.2221 0.6830 -0.33 0.75
## Year2005 -0.3543 0.6664 -0.53 0.60
## Year2006 -0.0714 0.7201 -0.10 0.92
```

```

## Year2007          -0.3951      0.6810   -0.58      0.56
## Year2008          -0.4424      0.6649   -0.67      0.51
## Year2009          -0.8921      0.6811   -1.31      0.19
## Year2010          -0.3330      0.6632   -0.50      0.62
## Year2011          -0.8442      0.6807   -1.24      0.22
## Year2012          -0.0913      0.6728   -0.14      0.89
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.306, Adjusted R-squared:  0.121
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.466  0.911  0.963  0.924  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      1.22e-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 3.903 1          1.976
## Year            3.903 16          1.043

```

Residuals from last author



```
## [1] "List of 0 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
## -1.1215 -0.2251 0.0177 0.2319 0.7456
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5531 0.6741 2.30 0.024 *
## LastAuthorFemale1 0.1534 0.1288 1.19 0.238
## Year1997 -0.4316 0.7536 -0.57 0.569
## Year1998 -0.8636 0.7615 -1.13 0.261
## Year1999 -0.2951 0.6741 -0.44 0.663
## Year2000 -0.0906 0.7180 -0.13 0.900
## Year2001 -0.3189 0.7801 -0.41 0.684
## Year2002 -0.4203 0.7673 -0.55 0.586
## Year2003 -0.3853 0.7204 -0.53 0.595
## Year2004 -0.3135 0.7416 -0.42 0.674
## Year2005 -0.3875 0.6950 -0.56 0.579
## Year2006 -0.1397 0.7816 -0.18 0.859
```

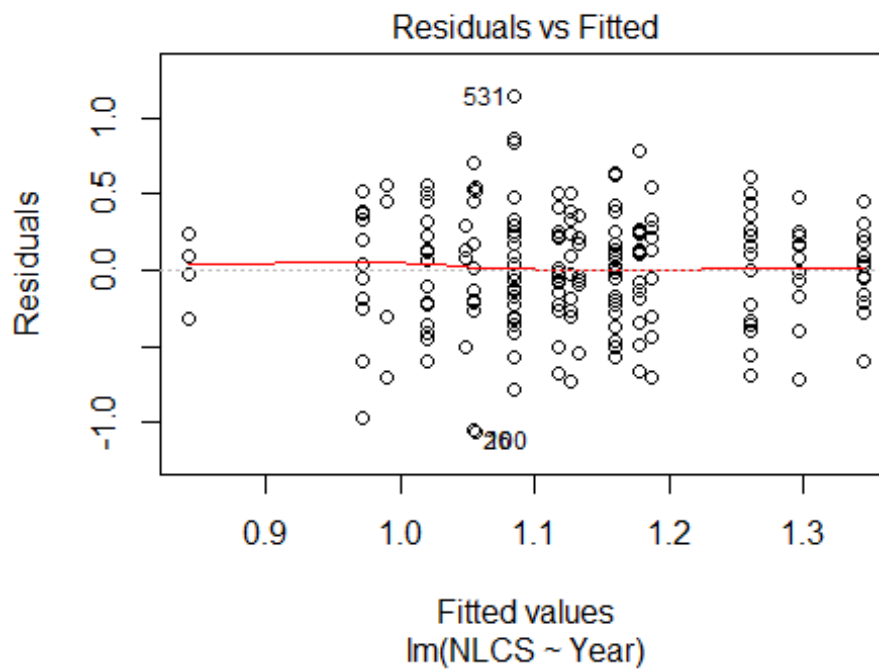


```

## Year2007          -0.4644      0.7206   -0.64    0.522
## Year2008          -0.4770      0.6925   -0.69    0.493
## Year2009          -0.9202      0.7041   -1.31    0.196
## Year2010          -0.3553      0.6863   -0.52    0.606
## Year2011          -0.8787      0.7042   -1.25    0.217
## Year2012          -0.1156      0.6911   -0.17    0.868
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.428
## Multiple R-squared:  0.329, Adjusted R-squared:  0.151
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 72 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.472  0.897   0.965   0.918   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      1.22e-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: 82"
## [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
##   20   11   12   12   13   13   14   17   13   24   22   22   28   32   31
## 2011 2012
##   36   48
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10    4    4    4    7    3    9   13   11   13   12   11   18   18   15
## 2011 2012

```

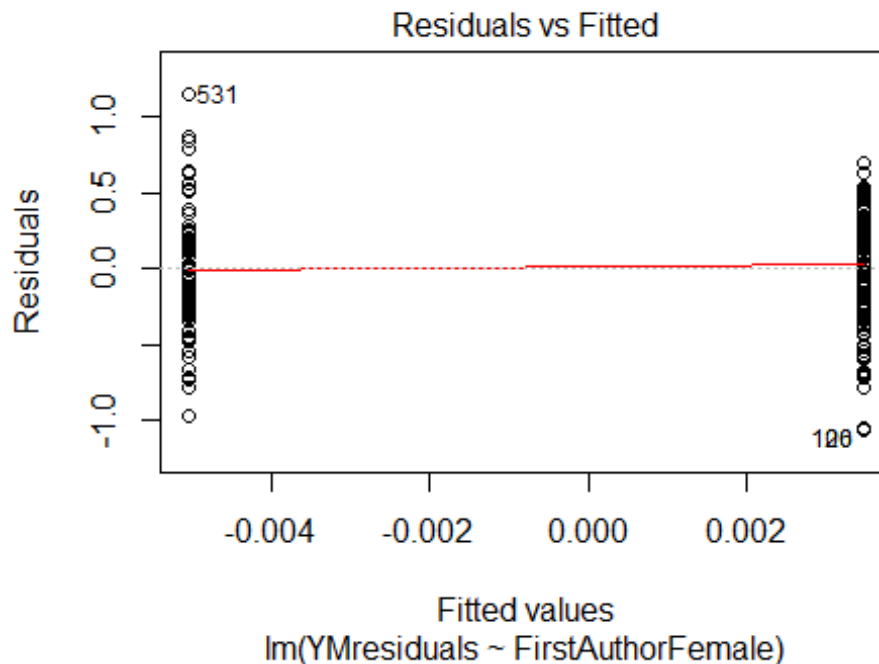
```
## 21 31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 4 4 3 5 3 7 13 10 9 10 7 16 13 15
## 2011 2012
## 17 24
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 16, df = 16, p-value = 0.4
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.41, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 5.18043572051951"
## [1] "Male first author team size 2018 geometric mean: 4.37004699438569"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

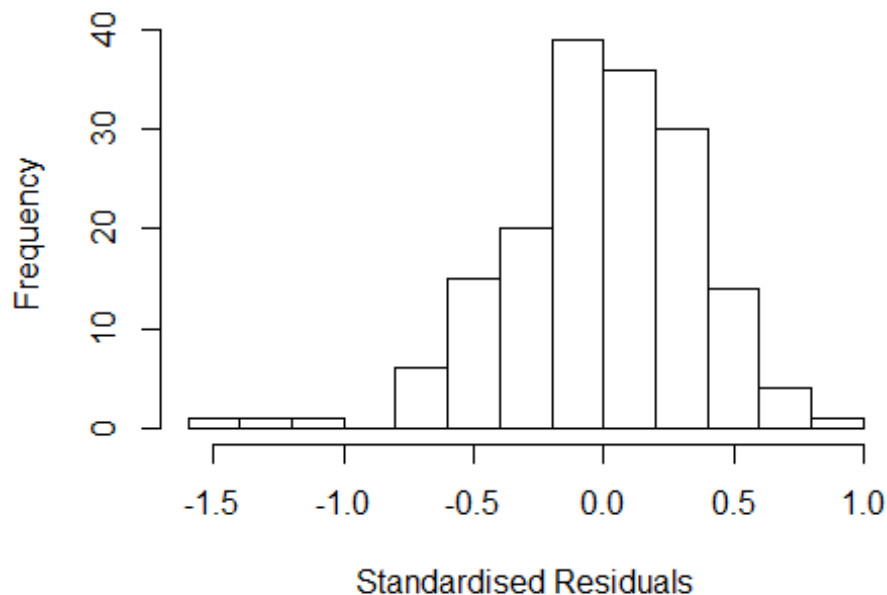
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 77, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6.32104156640794"
## [1] "Male last author team size 2018 geometric mean: 4.26954302508486"

## 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 = 74, 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 2.006 1      1.416
## LastAuthorFemale  1.855 1      1.362
## UniqueAuthors    9.027 4      1.317
## Year             20.689 16     1.099
```

Residuals from first and last author and team size



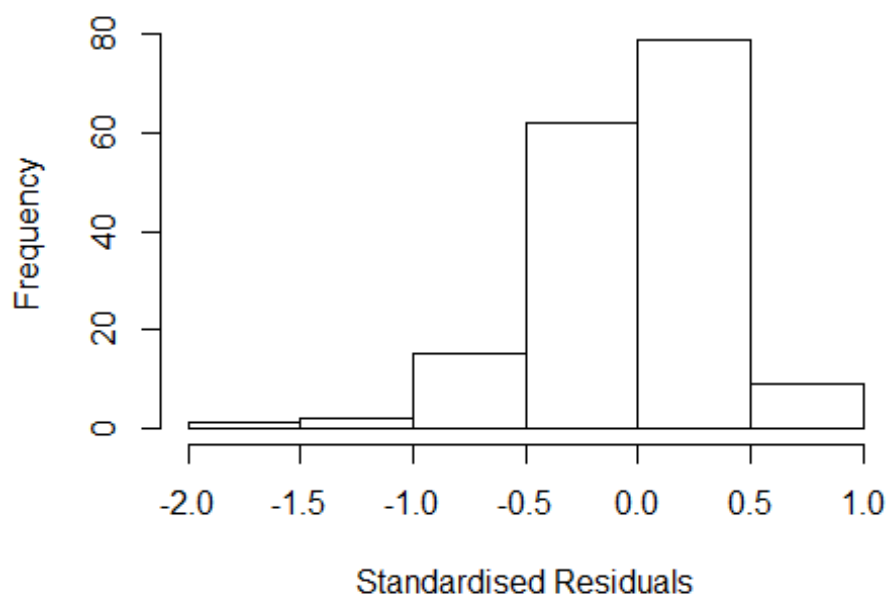
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.47185 -0.21071  0.00518  0.23822  0.90130
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.12402    0.19277    5.83 3.4e-08 ***
## FirstAuthorFemale1 -0.03344    0.06955   -0.48   0.63
## LastAuthorFemale1  0.01016    0.07319    0.14   0.89
## UniqueAuthors2     0.13631    0.18222    0.75   0.46
## UniqueAuthors3     0.07570    0.19603    0.39   0.70
## UniqueAuthors4     0.21576    0.17965    1.20   0.23
## UniqueAuthors5     0.21321    0.18196    1.17   0.24
## Year1997          -0.16197    0.47582   -0.34   0.73
## Year1998          -0.15933    0.25726   -0.62   0.54
## Year1999          -0.36846    0.22379   -1.65   0.10
```

```

## Year2000      -0.02742    0.24001   -0.11    0.91
## Year2001      0.27213    0.22201    1.23    0.22
## Year2002     -0.04756    0.26583   -0.18    0.86
## Year2003      0.05096    0.23396    0.22    0.83
## Year2004      0.00909    0.23608    0.04    0.97
## Year2005     -0.16617    0.24030   -0.69    0.49
## Year2006     -0.22779    0.26362   -0.86    0.39
## Year2007     -0.23015    0.22995   -1.00    0.32
## Year2008     -0.03200    0.24366   -0.13    0.90
## Year2009     -0.15754    0.22266   -0.71    0.48
## Year2010     -0.20442    0.24320   -0.84    0.40
## Year2011     -0.17655    0.22999   -0.77    0.44
## Year2012     -0.25609    0.22037   -1.16    0.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.152, Adjusted R-squared:  0.0236
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 153 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.019  0.861  0.948  0.898  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      5.95e-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 1.958 1      1.399
## LastAuthorFemale  1.587 1      1.260
## Year              3.044 16      1.035

```

Residuals from first and last author



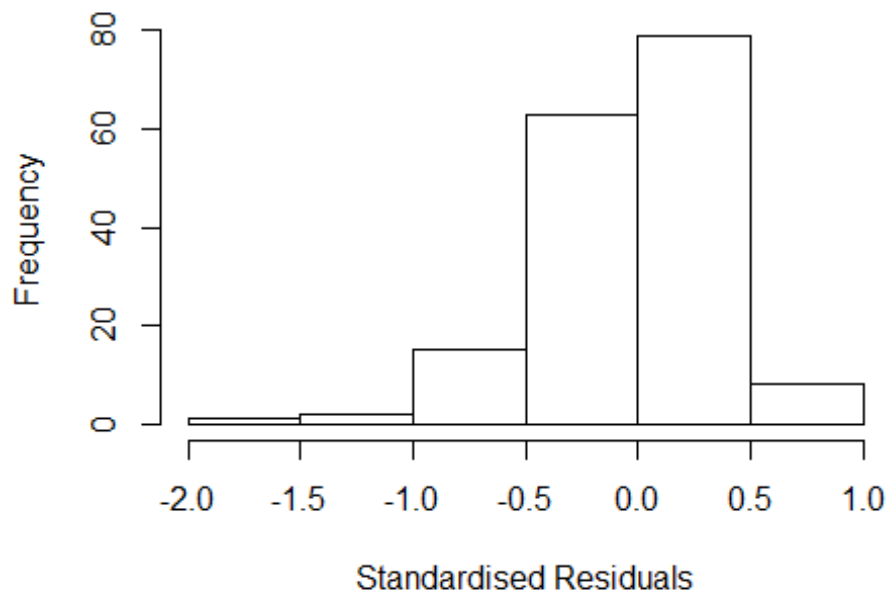
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5692 -0.2327 0.0094 0.2269 0.9628
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1896 0.1879 6.33 2.7e-09 ***
## FirstAuthorFemale1 -0.0133 0.0687 -0.19 0.847
## LastAuthorFemale1 0.0357 0.0675 0.53 0.598
## Year1997 -0.1583 0.4596 -0.34 0.731
## Year1998 -0.1294 0.2381 -0.54 0.587
## Year1999 -0.2547 0.1983 -1.28 0.201
## Year2000 0.0169 0.2066 0.08 0.935
## Year2001 0.3797 0.1864 2.04 0.043 *
## Year2002 0.0234 0.2491 0.09 0.925
## Year2003 0.1633 0.1980 0.82 0.411
## Year2004 0.1071 0.2140 0.50 0.618
## Year2005 -0.0729 0.2197 -0.33 0.740
```

```

## Year2006          -0.1645      0.2273   -0.72    0.470
## Year2007          -0.1194      0.2031   -0.59    0.558
## Year2008           0.0681      0.2088    0.33    0.745
## Year2009          -0.0726      0.2022   -0.36    0.720
## Year2010          -0.1777      0.2137   -0.83    0.407
## Year2011          -0.0887      0.2000   -0.44    0.658
## Year2012          -0.1901      0.1969   -0.97    0.336
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.362
## Multiple R-squared:  0.114, Adjusted R-squared:  0.00707
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 149 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0201 0.8730 0.9550 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      5.95e-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 2.241 1      1.497
## Year              2.241 16      1.026

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.5738 -0.2267 0.0127 0.2291 0.9539
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.20280 0.18457 6.52 1e-09 ***
## FirstAuthorFemale1 -0.01249 0.06843 -0.18 0.855
## Year1997 -0.17087 0.46296 -0.37 0.713
## Year1998 -0.13277 0.24152 -0.55 0.583
## Year1999 -0.25621 0.19492 -1.31 0.191
## Year2000 0.00352 0.20370 0.02 0.986
## Year2001 0.37097 0.19021 1.95 0.053 .
## Year2002 0.01784 0.25172 0.07 0.944
## Year2003 0.15307 0.19780 0.77 0.440
## Year2004 0.09777 0.21477 0.46 0.650
## Year2005 -0.08153 0.21230 -0.38 0.702
## Year2006 -0.16135 0.22800 -0.71 0.480
```



```

## Year2007          -0.12210      0.20263    -0.60      0.548
## Year2008           0.06369      0.20977      0.30      0.762
## Year2009          -0.08290      0.20209    -0.41      0.682
## Year2010          -0.17384      0.21357    -0.81      0.417
## Year2011          -0.08910      0.20088    -0.44      0.658
## Year2012          -0.19523      0.19714    -0.99      0.324
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.114, Adjusted R-squared:  0.0135
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 150 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0143 0.8740 0.9530 0.9030 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      5.95e-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 1.532 1      1.238
## Year      1.532 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.56997 -0.23309  0.00731  0.23083  0.95579

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1878    0.1907    6.23 4.5e-09 ***
## LastAuthorFemale1  0.0352    0.0681    0.52   0.61
## Year1997         -0.1649    0.4557   -0.36   0.72
## Year1998         -0.1275    0.2404   -0.53   0.60
## Year1999         -0.2528    0.2010   -1.26   0.21
## Year2000          0.0163    0.2088    0.08   0.94
## Year2001          0.3822    0.1842    2.07   0.04 *
## Year2002          0.0203    0.2505    0.08   0.94
## Year2003          0.1597    0.2024    0.79   0.43
## Year2004          0.0977    0.2133    0.46   0.65
## Year2005         -0.0758    0.2217   -0.34   0.73
## Year2006         -0.1670    0.2312   -0.72   0.47
## Year2007         -0.1245    0.2073   -0.60   0.55
## Year2008          0.0669    0.2117    0.32   0.75
## Year2009         -0.0765    0.2052   -0.37   0.71
## Year2010         -0.1853    0.2149   -0.86   0.39
## Year2011         -0.0914    0.2052   -0.45   0.66
## Year2012         -0.1946    0.2020   -0.96   0.34
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.361
## Multiple R-squared:  0.114, Adjusted R-squared:  0.0134
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 149 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0191 0.8710 0.9530 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      5.95e-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: 168"
## [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
##    2    7    2    4   26   12   14   13   10   21   19   11   17   19   12
## 2011 2012
##   18   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    1    1   12    3    9   11    7   13   14    6   11   14   10
## 2011 2012
##   13   16
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    1    1   11    3    7   11    7    9   11    4   11   13    9
## 2011 2012
##   11   16
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.30934976510868"
## [1] "Male first author team size 2018 geometric mean: 4.40429626701948"

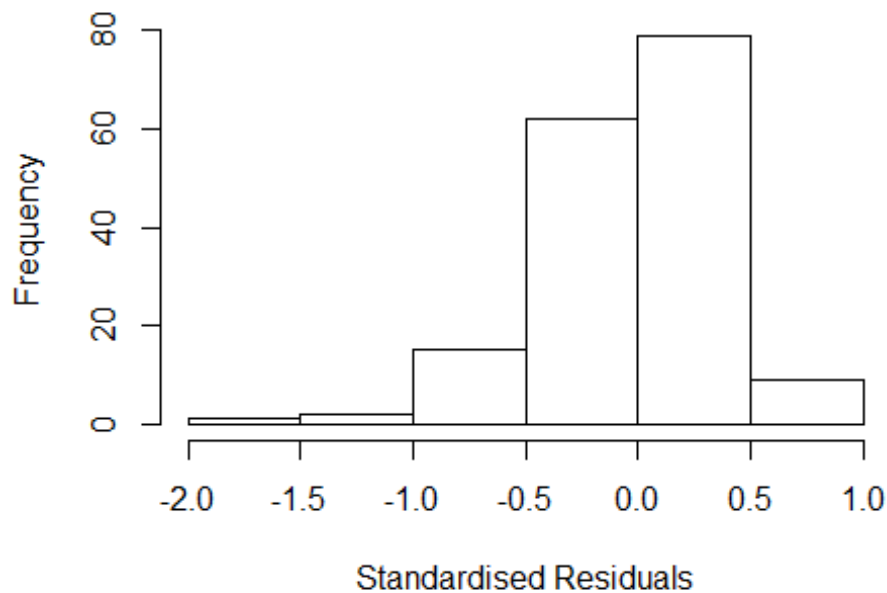
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 40, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.54666402862926"
## [1] "Male last author team size 2018 geometric mean: 4.39265951166886"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36, 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 3.065e+13  1      5.536e+06
## LastAuthorFemale  1.762e+13  1      4.198e+06
## UniqueAuthors    3.636e+14  4      6.608e+01
## Year              9.390e+27 16      7.484e+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 +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
```

```

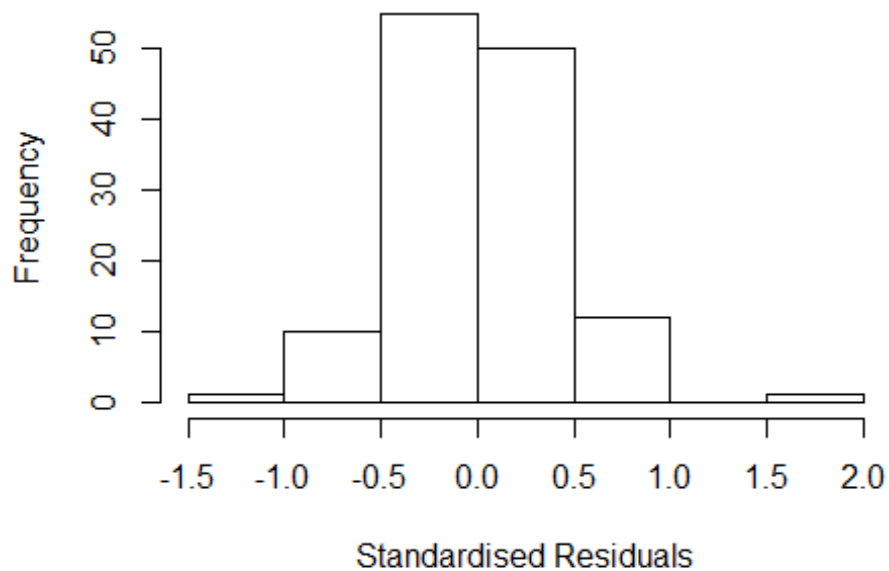
##           Min           1Q           Median           3Q           Max
## -1.08e+00 -2.79e-01  8.67e-16  2.72e-01  1.59e+00
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4252     0.1900   7.50 2.0e-11 ***
## FirstAuthorFemale1 -0.0467     0.0883  -0.53 0.59841
## LastAuthorFemale1 -0.0930     0.1083  -0.86 0.39268
## UniqueAuthors2     0.3402     0.2288   1.49 0.13999
## UniqueAuthors3     0.2552     0.1889   1.35 0.17973
## UniqueAuthors4     0.3398     0.1900   1.79 0.07649 .
## UniqueAuthors5     0.3488     0.1930   1.81 0.07349 .
## Year1997          -1.0966     0.1916  -5.72 9.9e-08 ***
## Year1998          -0.8023     0.1183  -6.78 6.9e-10 ***
## Year1999          -0.0214     0.1116  -0.19 0.84869
## Year2000          -1.2711     0.2380  -5.34 5.4e-07 ***
## Year2001          -1.1229     0.6603  -1.70 0.09196 .
## Year2002          -0.5689     0.2010  -2.83 0.00557 **
## Year2003          -0.3564     0.1107  -3.22 0.00171 **
## Year2004          -0.8189     0.1980  -4.14 7.1e-05 ***
## Year2005          -0.5967     0.1508  -3.96 0.00014 ***
## Year2006          -0.6433     0.1419  -4.53 1.5e-05 ***
## Year2007          -0.9281     0.4976  -1.87 0.06491 .
## Year2008          -0.5442     0.1810  -3.01 0.00329 **
## Year2009          -0.7724     0.1032  -7.48 2.2e-11 ***
## Year2010          -0.5301     0.1847  -2.87 0.00495 **
## Year2011          -0.5145     0.1614  -3.19 0.00188 **
## Year2012          -0.6866     0.1157  -5.94 3.7e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.371
## Multiple R-squared:  0.38,   Adjusted R-squared:  0.251
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 116 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0272 0.8790 0.9400 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           7.75e-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"

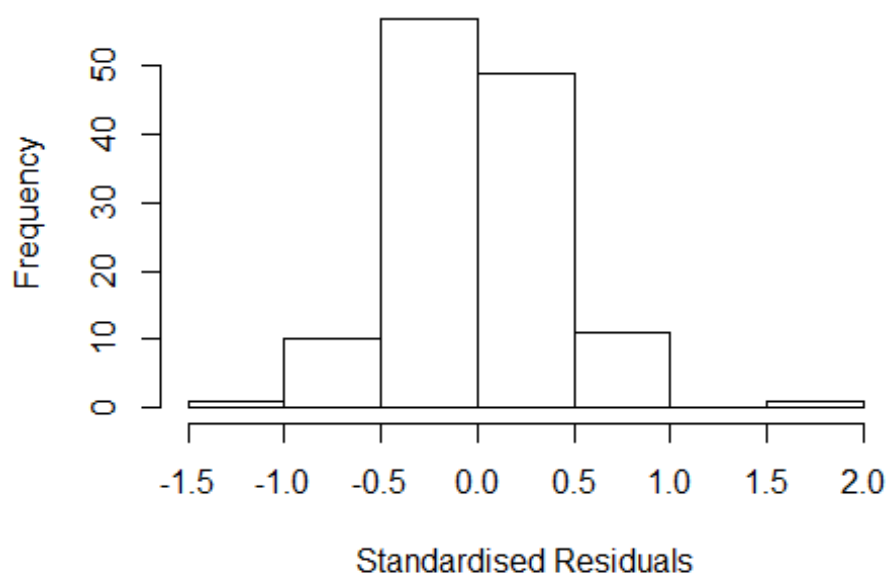
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead
```

Residuals from first and last author and team size



```
##               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 2.044e+14  1      1.430e+07
## LastAuthorFemale  2.087e+15  1      4.569e+07
## Year              2.186e+28 16      7.684e+00
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.10037 -0.26037 -0.00672 0.24863 1.80706
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.76500 0.00000 Inf < 2e-16 ***
## FirstAuthorFemale1 0.00943 0.07445 0.13 0.89946
## LastAuthorFemale1 -0.11578 0.10188 -1.14 0.25826
## Year1997 -1.20062 0.11859 -10.12 < 2e-16 ***
## Year1998 -0.84943 0.07445 -11.41 < 2e-16 ***
## Year1999 -0.05465 0.12254 -0.45 0.65646
## Year2000 -1.45928 0.18987 -7.69 6.8e-12 ***
## Year2001 -1.18081 0.72608 -1.63 0.10675
## Year2002 -0.59322 0.18735 -3.17 0.00200 **
## Year2003 -0.42836 0.06469 -6.62 1.3e-09 ***
## Year2004 -0.89054 0.17717 -5.03 2.0e-06 ***
## Year2005 -0.65580 0.09429 -6.96 2.6e-10 ***
```

```

## Year2006          -0.67406      0.12168      -5.54      2.1e-07 ***
## Year2007          -0.91057      0.56409      -1.61      0.10934
## Year2008          -0.55361      0.17201      -3.22      0.00169 **
## Year2009          -0.83172      0.10100      -8.23      4.1e-13 ***
## Year2010          -0.59946      0.15648      -3.83      0.00021 ***
## Year2011          -0.51658      0.13961      -3.70      0.00034 ***
## Year2012          -0.79428      0.08696      -9.13      3.8e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.372, Adjusted R-squared:  0.269
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## observation 13 is an outlier with |weight| = 0 ( < 0.00078);
## 13 weights are ~= 1. The remaining 115 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.323  0.868  0.942  0.897  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          7.75e-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 -1.422e+15  1          NaN
## Year              -1.422e+15 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 = "MM"
## Residuals:

```



```

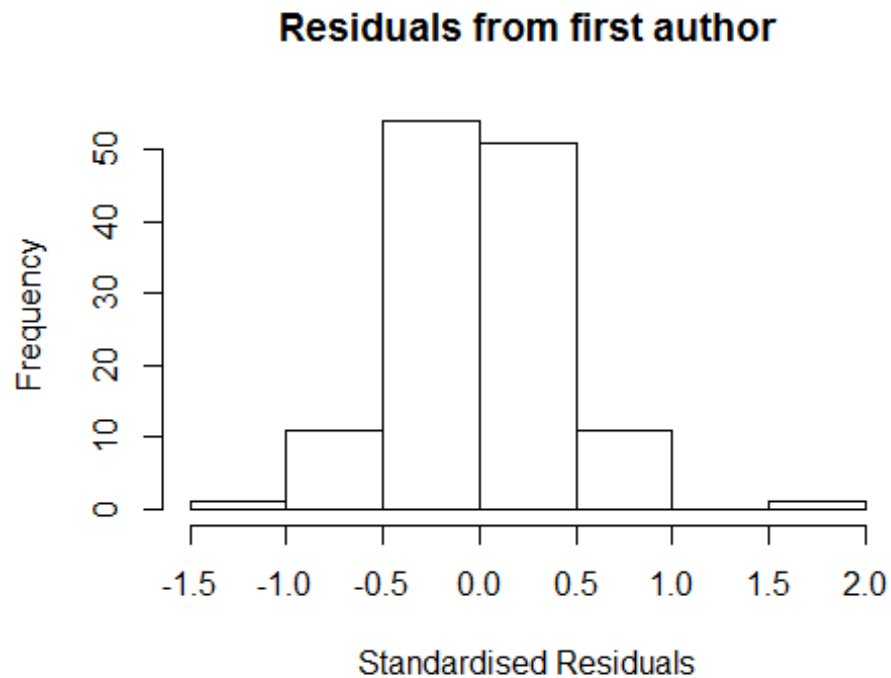
##           Min           1Q           Median           3Q           Max
## -1.08e+00 -2.74e-01  2.22e-16  2.71e-01  1.71e+00
##
## Coefficients:
##               Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.76e+00  4.91e-08  3.60e+07 < 2e-16 ***
## FirstAuthorFemale1 1.38e-03  7.60e-02  2.00e-02  0.98553
## Year1997        -1.20e+00  1.19e-01 -1.00e+01 < 2e-16 ***
## Year1998        -8.41e-01  7.60e-02 -1.11e+01 < 2e-16 ***
## Year1999        -1.62e-01  7.60e-02 -2.14e+00  0.03481 *
## Year2000        -1.48e+00  1.91e-01 -7.74e+00  5.0e-12 ***
## Year2001        -1.16e+00  6.83e-01 -1.70e+00  0.09173 .
## Year2002        -6.10e-01  1.74e-01 -3.50e+00  0.00067 ***
## Year2003        -4.37e-01  6.37e-02 -6.86e+00  4.0e-10 ***
## Year2004        -9.44e-01  1.77e-01 -5.33e+00  5.3e-07 ***
## Year2005        -6.80e-01  9.31e-02 -7.31e+00  4.5e-11 ***
## Year2006        -6.89e-01  1.18e-01 -5.85e+00  5.0e-08 ***
## Year2007        -9.70e-01  4.63e-01 -2.10e+00  0.03840 *
## Year2008        -5.65e-01  1.62e-01 -3.50e+00  0.00068 ***
## Year2009        -8.28e-01  1.01e-01 -8.22e+00  4.1e-13 ***
## Year2010        -6.05e-01  1.59e-01 -3.82e+00  0.00022 ***
## Year2011        -5.78e-01  1.22e-01 -4.75e+00  6.2e-06 ***
## Year2012        -8.11e-01  8.93e-02 -9.08e+00  4.7e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.367
## Multiple R-squared:  0.359, Adjusted R-squared:  0.261
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## observation 13 is an outlier with |weight| <= 5.6e-05 ( < 0.00078);
## 16 weights are ~= 1. The remaining 112 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.369  0.875   0.946   0.898   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           7.75e-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"

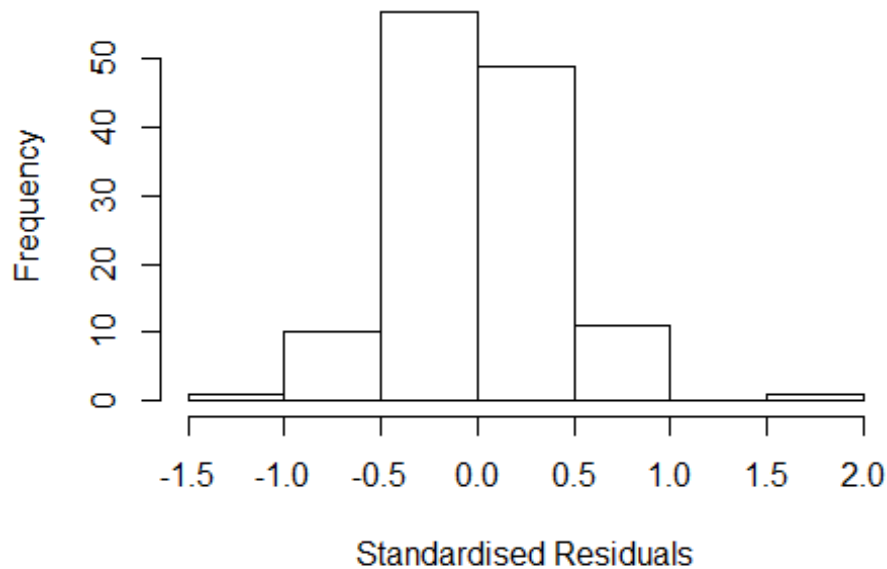
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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 16          NaN
```

Residuals from last author



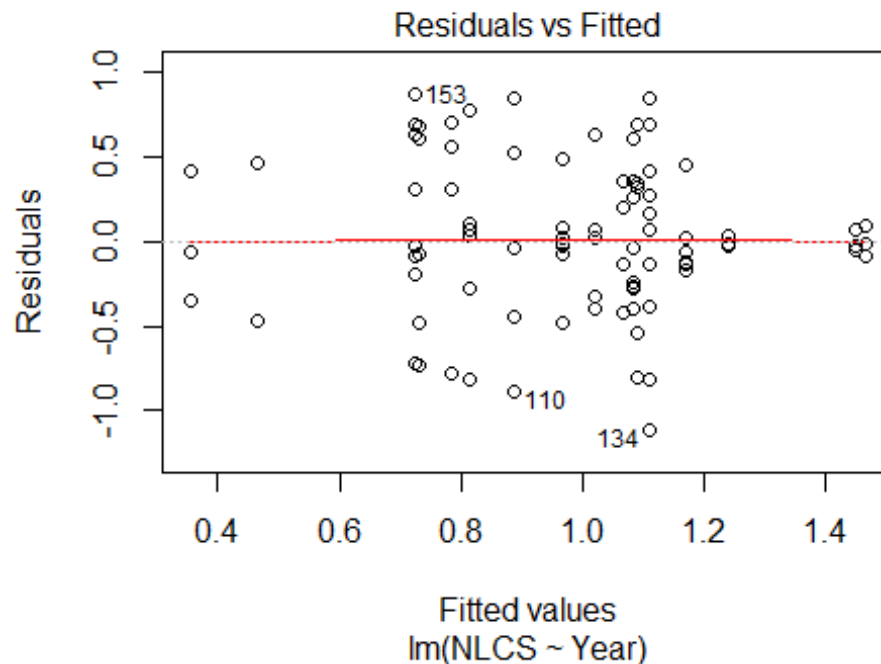
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.09763 -0.25763 -0.00526 0.24937 1.80377
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7650 0.0000 Inf < 2e-16 ***
## LastAuthorFemale1 -0.1156 0.1024 -1.13 0.26137
## Year1997 -1.1972 0.1159 -10.33 < 2e-16 ***
## Year1998 -0.8400 0.0000 -Inf < 2e-16 ***
## Year1999 -0.0454 0.1024 -0.44 0.65813
## Year2000 -1.4562 0.1786 -8.15 5.9e-13 ***
## Year2001 -1.1954 0.7450 -1.60 0.11141
## Year2002 -0.5886 0.1876 -3.14 0.00218 **
## Year2003 -0.4260 0.0610 -6.98 2.3e-10 ***
## Year2004 -0.8870 0.1696 -5.23 8.1e-07 ***
## Year2005 -0.6507 0.0856 -7.61 9.9e-12 ***
## Year2006 -0.6674 0.1222 -5.46 2.9e-07 ***
```

```

## Year2007          -0.9050      0.5965    -1.52   0.13208
## Year2008          -0.5504      0.1755    -3.14   0.00220 **
## Year2009          -0.8269      0.0909    -9.10   4.2e-15 ***
## Year2010          -0.5962      0.1582    -3.77   0.00027 ***
## Year2011          -0.5104      0.1374    -3.72   0.00032 ***
## Year2012          -0.7892      0.0795    -9.93   < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.351
## Multiple R-squared:  0.376, Adjusted R-squared:  0.28
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## observation 13 is an outlier with |weight| = 0 ( < 0.00078);
## 12 weights are ~= 1. The remaining 116 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.308  0.867  0.940  0.894  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      7.75e-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: 129"
## [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
##    9   11   12    6    3    6    5    8    5    4    9    9    8    4   11
## 2011 2012
##    8   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    5    7    4    3    3    5    7    5    2    8    6    5    3   10

```

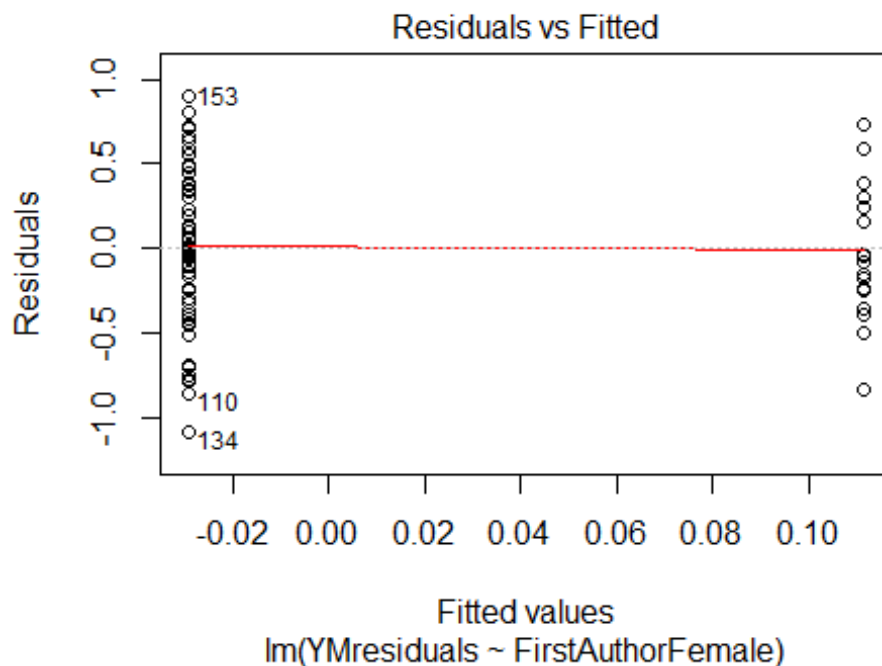
```
## 2011 2012
## 5 10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 3 5 4 3 3 4 4 4 2 7 6 4 3 10
## 2011 2012
## 4 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 0.35, df = 1, p-value = 0.6
## [1] "Female first author team size 2018 geometric mean: 6"
## [1] "Male first author team size 2018 geometric mean: 3"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
```

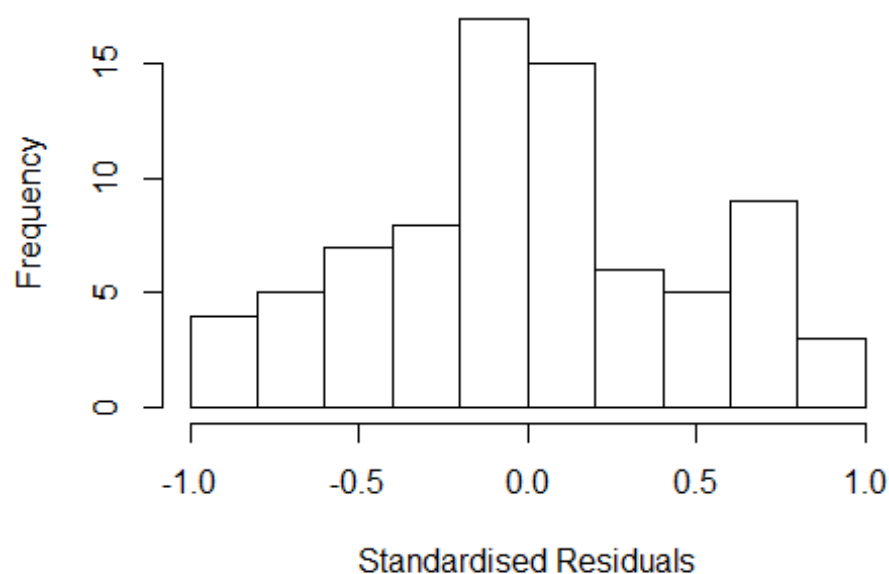
```
## W = 1, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6"
## [1] "Male last author team size 2018 geometric mean: 3"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, 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"

## Warning in lmrob.fit(x, y, control, init = init, mf = mf): M-step did NOT
## converge. Returning unconverged SM-estimate
```



```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.891 1      1.973
## LastAuthorFemale  5.489 1      2.343
## Year             20.846 16      1.100
```

Residuals from first and last author



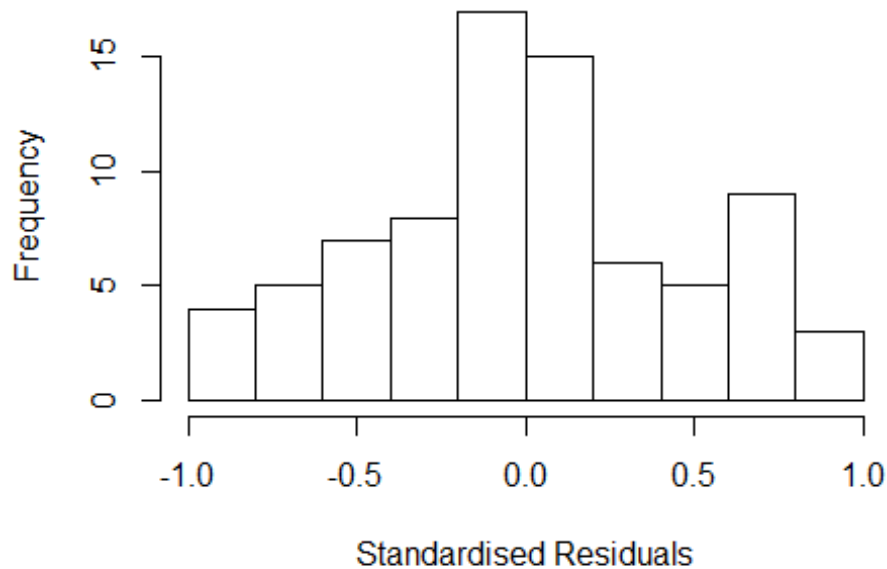
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9936 -0.3064 -0.0113 0.3160 0.8857
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.35110 0.19144 1.83 0.0716 .
## FirstAuthorFemale1 0.26549 0.20714 1.28 0.2049
## LastAuthorFemale1 -0.00146 0.18750 -0.01 0.9938
## Year1997 0.37251 0.36706 1.01 0.3143
## Year1998 0.53434 0.30305 1.76 0.0830 .
## Year1999 0.71768 0.24741 2.90 0.0052 **
## Year2000 0.88890 0.19203 4.63 2.0e-05 ***
## Year2001 1.11421 0.19637 5.67 4.3e-07 ***
## Year2002 0.19515 0.36961 0.53 0.5995
## Year2003 0.53017 0.24242 2.19 0.0327 *
## Year2004 0.64973 0.29124 2.23 0.0294 *
## Year2005 0.11240 0.42728 0.26 0.7934
```

```

## Year2006          0.64882      0.25661      2.53      0.0141 *
## Year2007          0.67699      0.25174      2.69      0.0093 **
## Year2008          0.50003      0.45640      1.10      0.2776
## Year2009          1.09803      0.20347      5.40      1.2e-06 ***
## Year2010          0.64254      0.34344      1.87      0.0662 .
## Year2011          0.28794      0.49435      0.58      0.5624
## Year2012          0.35616      0.30310      1.18      0.2446
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.535
## Multiple R-squared:  0.261, Adjusted R-squared:  0.0397
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 64 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##   0.711  0.865   0.948   0.920   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      1.27e-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 3.951 1      1.988
## Year              3.951 16      1.044

```


Residuals from first author



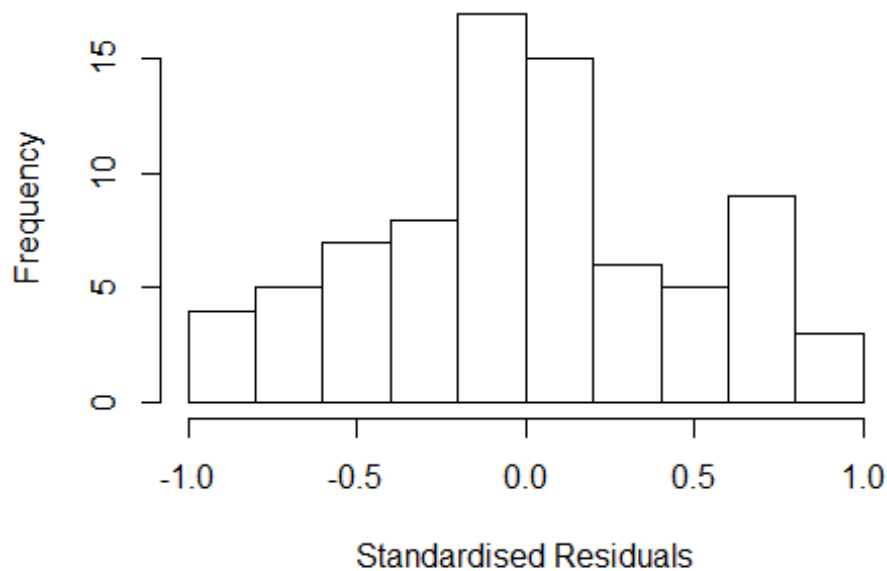
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9939 -0.3056 -0.0113 0.3159 0.8854
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.351 0.192 1.83 0.0723 .
## FirstAuthorFemale1 0.266 0.209 1.27 0.2094
## Year1997 0.371 0.369 1.01 0.3175
## Year1998 0.535 0.303 1.76 0.0826 .
## Year1999 0.718 0.248 2.90 0.0052 **
## Year2000 0.889 0.192 4.62 2.0e-05 ***
## Year2001 1.114 0.197 5.66 4.3e-07 ***
## Year2002 0.193 0.358 0.54 0.5916
## Year2003 0.530 0.236 2.25 0.0282 *
## Year2004 0.649 0.292 2.22 0.0299 *
## Year2005 0.113 0.430 0.26 0.7942
## Year2006 0.648 0.255 2.54 0.0137 *
```

```

## Year2007          0.677      0.253      2.68      0.0095 **
## Year2008          0.501      0.466      1.08      0.2866
## Year2009          1.098      0.194      5.66      4.4e-07 ***
## Year2010          0.643      0.335      1.92      0.0593 .
## Year2011          0.288      0.497      0.58      0.5650
## Year2012          0.357      0.305      1.17      0.2464
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.524
## Multiple R-squared:  0.263, Adjusted R-squared:  0.0571
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 65 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.699  0.862  0.947   0.918  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      1.27e-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 5.494 1      2.344
## Year              5.494 16      1.055

```

Residuals from last author



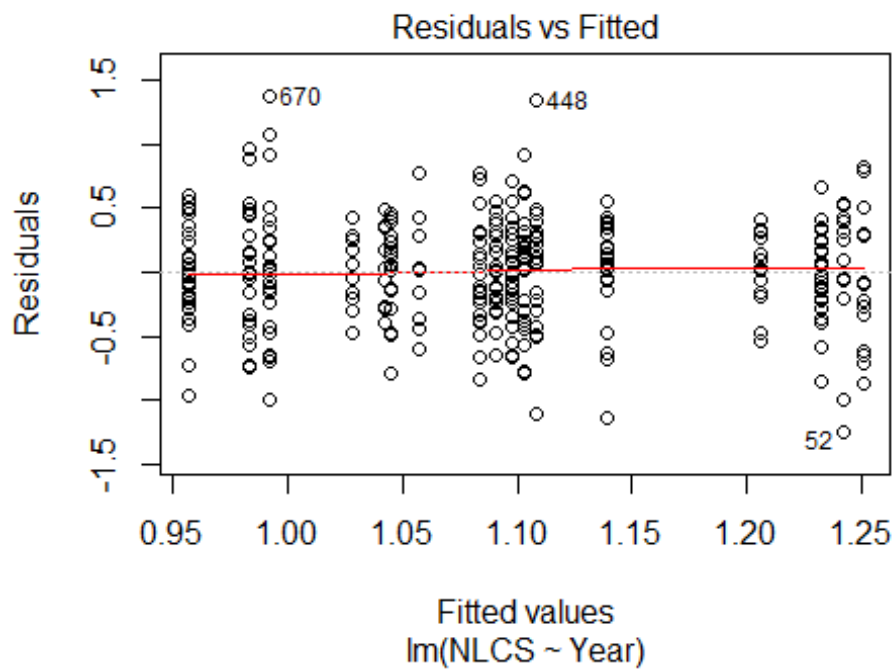
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1384 -0.3103 -0.0223 0.3327 0.8772
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3510 0.1918 1.83 0.07220 .
## LastAuthorFemale1 0.0305 0.2046 0.15 0.88190
## Year1997 0.3716 0.3686 1.01 0.31738
## Year1998 0.5353 0.3036 1.76 0.08282 .
## Year1999 0.7107 0.2521 2.82 0.00650 **
## Year2000 0.8890 0.1924 4.62 2.0e-05 ***
## Year2001 1.1143 0.1968 5.66 4.3e-07 ***
## Year2002 0.1838 0.3754 0.49 0.62613
## Year2003 0.5818 0.2747 2.12 0.03826 *
## Year2004 0.6492 0.2919 2.22 0.02987 *
## Year2005 0.1125 0.4291 0.26 0.79403
## Year2006 0.7629 0.2379 3.21 0.00214 **
```

```

## Year2007          0.8147      0.2097      3.89  0.00025 ***
## Year2008          0.5561      0.5541      1.00  0.31954
## Year2009          1.0875      0.2063      5.27  1.9e-06 ***
## Year2010          0.7874      0.3111      2.53  0.01396 *
## Year2011          0.3460      0.5707      0.61  0.54654
## Year2012          0.3678      0.2895      1.27  0.20876
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.526
## Multiple R-squared:  0.243, Adjusted R-squared:  0.0326
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 64 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.618  0.850  0.937   0.915  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      1.27e-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: 79"
## [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
##   20   22   27   20   44   30   31   23   35   25   29   29   28   30   39
## 2011 2012
##   34   41
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   16   12   10   24   10   20   15   29   21   20   26   22   25   27
## 2011 2012

```

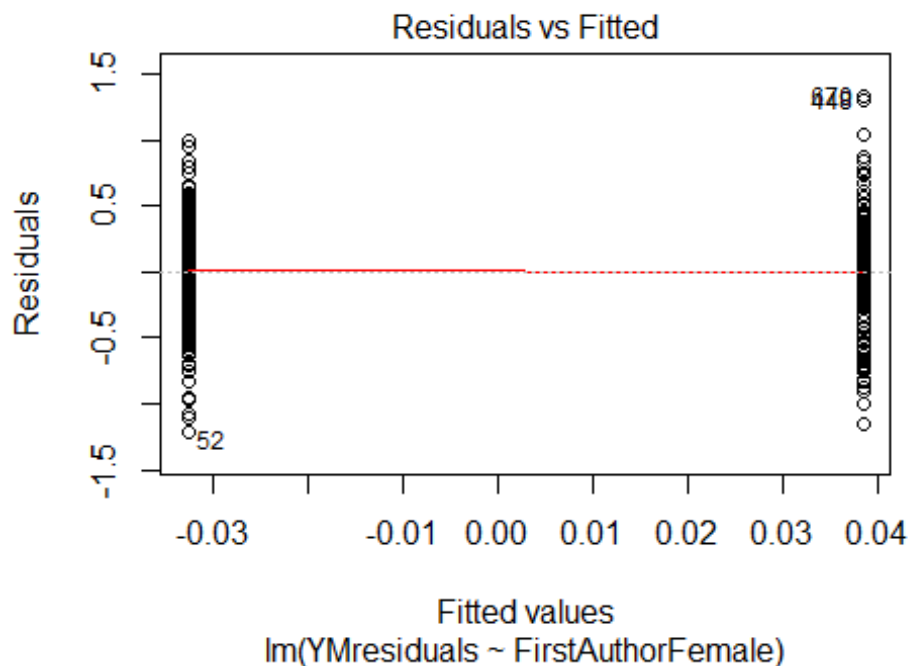
```
## 25 30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 14 10 8 22 8 16 11 24 18 18 25 20 23 25
## 2011 2012
## 23 28
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 0.17, df = 1, p-value = 0.7
## [1] "Female first author team size 2018 geometric mean: 4.20118321158612"
## [1] "Male first author team size 2018 geometric mean: 3.72640117233658"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

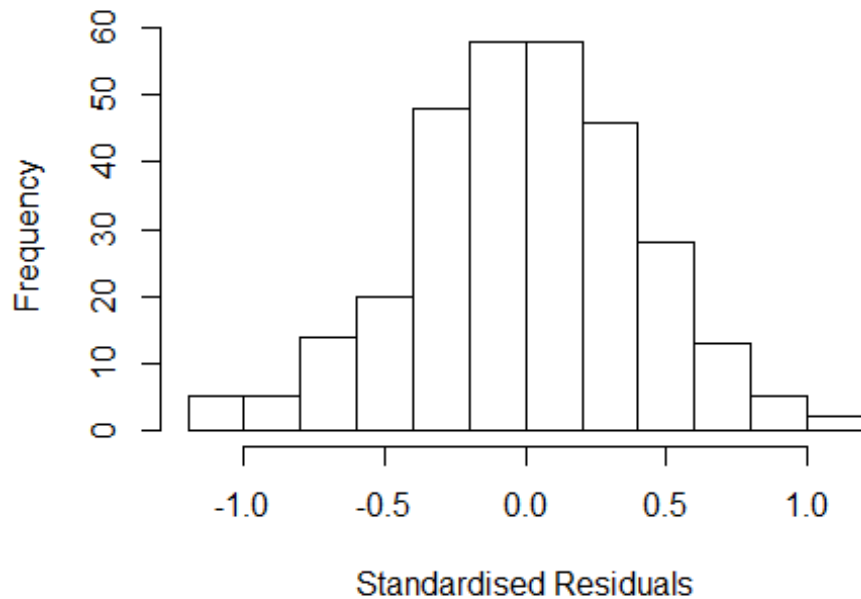
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.43093628992595"
## [1] "Male last author team size 2018 geometric mean: 4.14188718334238"

## 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 = 68, 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 1.392  1      1.180
## LastAuthorFemale  1.383  1      1.176
## UniqueAuthors    2.312  4      1.110
## Year              3.708 16      1.042
```

Residuals from first and last author and team size



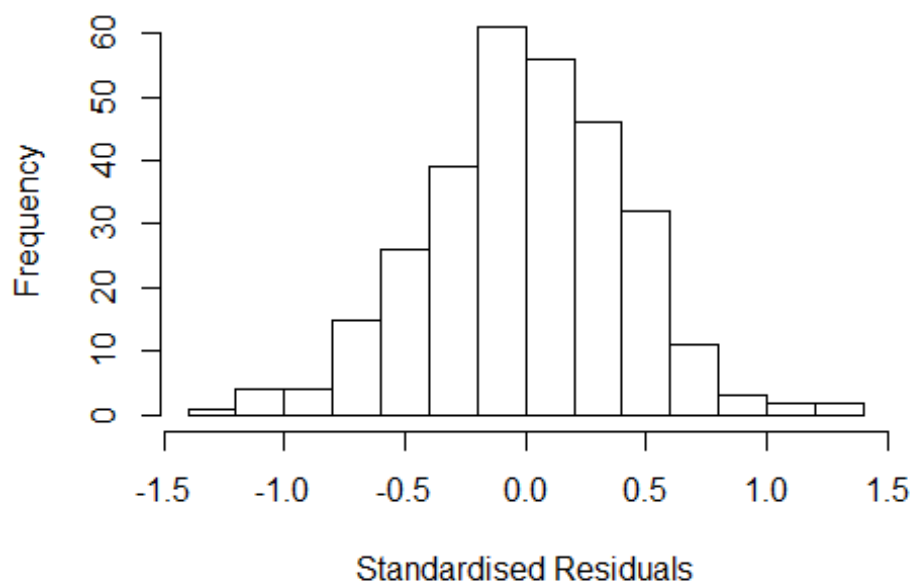
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.19706 -0.25876 0.00281 0.27667 1.18300
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.88404 0.13049 6.77 7.4e-11 ***
## FirstAuthorFemale1 0.08637 0.05286 1.63 0.10342
## LastAuthorFemale1 0.06122 0.05299 1.16 0.24895
## UniqueAuthors2 0.08208 0.09600 0.85 0.39330
## UniqueAuthors3 0.18619 0.09142 2.04 0.04263 *
## UniqueAuthors4 0.12556 0.09881 1.27 0.20487
## UniqueAuthors5 0.36262 0.10470 3.46 0.00062 ***
## Year1997 0.19112 0.22026 0.87 0.38630
## Year1998 0.15592 0.20269 0.77 0.44240
## Year1999 -0.08349 0.16560 -0.50 0.61453
```

```

## Year2000      -0.00926    0.14074   -0.07  0.94757
## Year2001      0.03332    0.19605    0.17  0.86518
## Year2002      0.15500    0.12642    1.23  0.22122
## Year2003      0.05558    0.15270    0.36  0.71613
## Year2004      0.16232    0.12289    1.32  0.18763
## Year2005     -0.05721    0.13176   -0.43  0.66450
## Year2006     -0.05060    0.12183   -0.42  0.67818
## Year2007     -0.16515    0.12846   -1.29  0.19965
## Year2008      0.04046    0.14943    0.27  0.78678
## Year2009     -0.09982    0.13784   -0.72  0.46957
## Year2010     -0.06660    0.12677   -0.53  0.59977
## Year2011      0.03189    0.14107    0.23  0.82132
## Year2012     -0.15103    0.14034   -1.08  0.28278
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.115, Adjusted R-squared:  0.0451
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 273 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.357  0.878  0.947  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      3.31e-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 1.354 1      1.164
## LastAuthorFemale  1.390 1      1.179
## Year              1.765 16      1.018

```


Residuals from first and last author



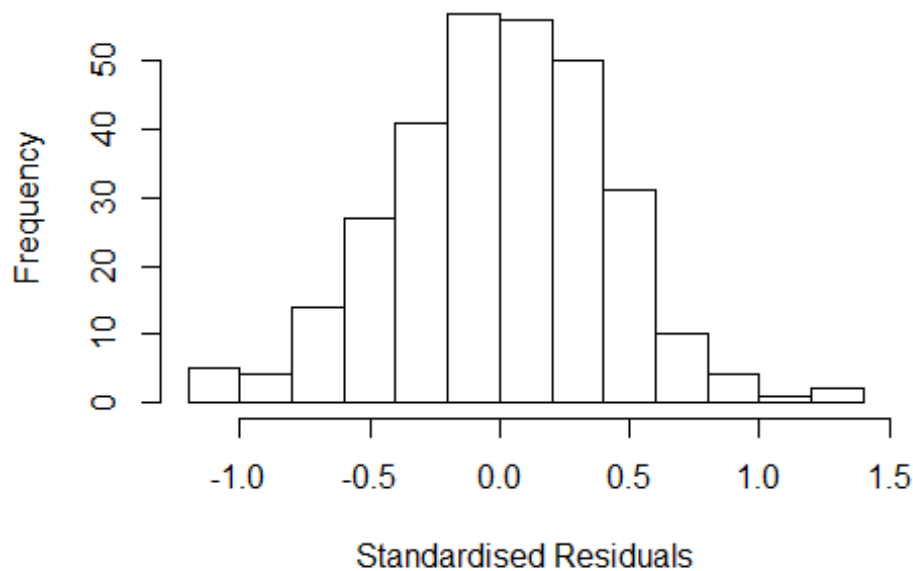
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.20514 -0.25700 0.00878 0.27757 1.38875
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9755 0.1028 9.49 <2e-16 ***
## FirstAuthorFemale1 0.0860 0.0534 1.61 0.109
## LastAuthorFemale1 0.0491 0.0528 0.93 0.353
## Year1997 0.2325 0.2423 0.96 0.338
## Year1998 0.1805 0.1963 0.92 0.359
## Year1999 -0.0102 0.1501 -0.07 0.946
## Year2000 0.0710 0.1321 0.54 0.592
## Year2001 0.0725 0.1898 0.38 0.703
## Year2002 0.1960 0.1192 1.64 0.101
## Year2003 0.1273 0.1467 0.87 0.386
## Year2004 0.2196 0.1192 1.84 0.066 .
## Year2005 0.0234 0.1213 0.19 0.847
```

```

## Year2006          0.0349      0.1174      0.30      0.767
## Year2007         -0.0506      0.1188     -0.43      0.670
## Year2008          0.1044      0.1514      0.69      0.491
## Year2009         -0.0257      0.1474     -0.17      0.862
## Year2010          0.0386      0.1253      0.31      0.758
## Year2011          0.0244      0.1324      0.18      0.854
## Year2012         -0.0853      0.1411     -0.60      0.546
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.417
## Multiple R-squared:  0.0606, Adjusted R-squared:  0.00085
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 276 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.245  0.877   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      3.31e-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 1.299 1          1.140
## Year              1.299 16          1.008

```

Residuals from first author



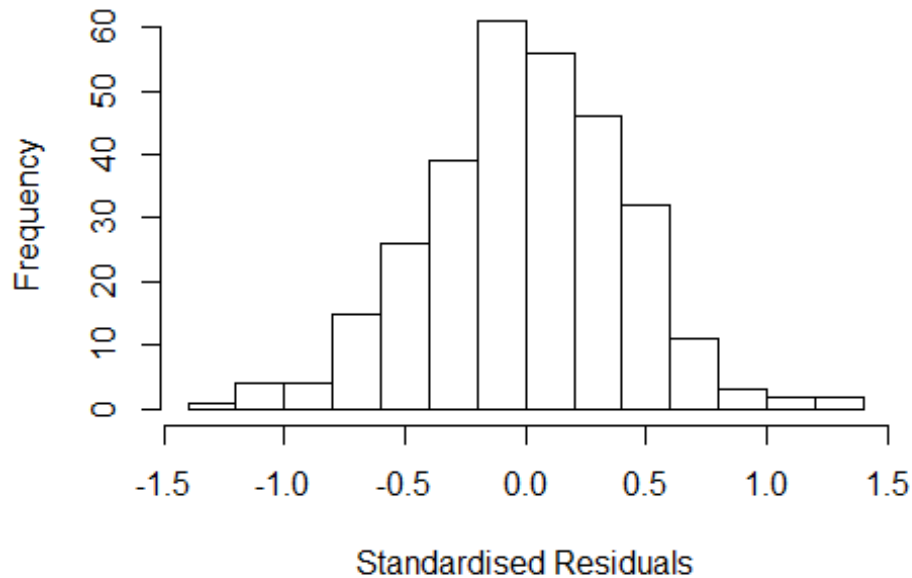
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1851 -0.25935 0.00554 0.28142 1.36718
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.99004 0.09980 9.92 <2e-16 ***
## FirstAuthorFemale1 0.09165 0.05271 1.74 0.083 .
## Year1997 0.23009 0.24312 0.95 0.345
## Year1998 0.18650 0.19210 0.97 0.332
## Year1999 -0.00713 0.14999 -0.05 0.962
## Year2000 0.06756 0.13187 0.51 0.609
## Year2001 0.08284 0.19259 0.43 0.667
## Year2002 0.18912 0.11910 1.59 0.113
## Year2003 0.12250 0.14715 0.83 0.406
## Year2004 0.21234 0.11668 1.82 0.070 .
## Year2005 0.01855 0.12061 0.15 0.878
## Year2006 0.03034 0.11717 0.26 0.796
```

```

## Year2007      -0.05192    0.11915   -0.44    0.663
## Year2008      0.10682    0.15021    0.71    0.478
## Year2009     -0.02592    0.14778   -0.18    0.861
## Year2010      0.04751    0.12477    0.38    0.704
## Year2011      0.03590    0.13132    0.27    0.785
## Year2012     -0.08387    0.13961   -0.60    0.548
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.417
## Multiple R-squared:  0.0574, Adjusted R-squared:  0.00101
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 276 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.260  0.879  0.955  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      3.31e-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 1.371 1      1.171
## Year      1.371 16      1.010

```

Residuals from last author



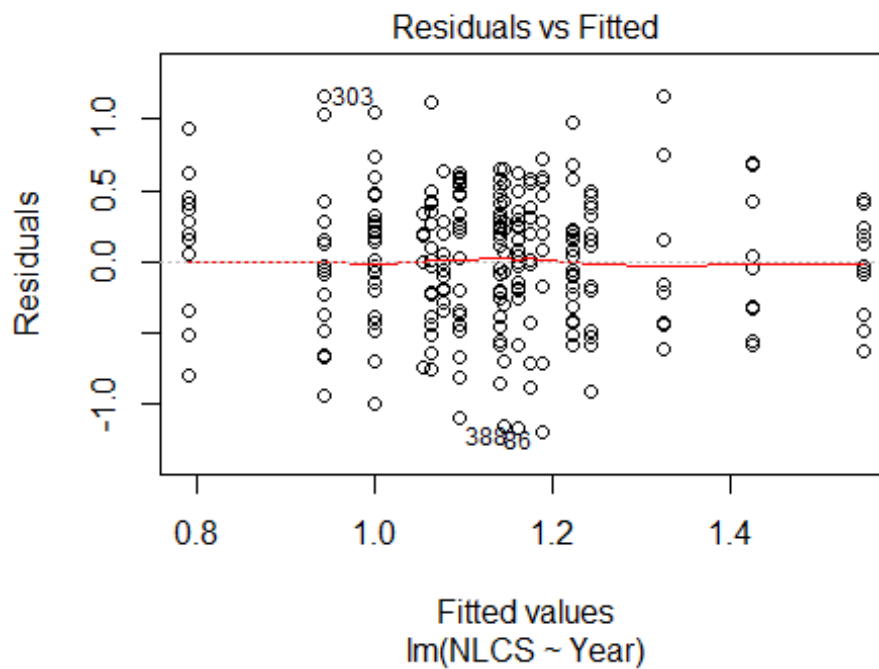
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.28807 -0.26776 0.00776 0.28654 1.43266
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02811 0.10135 10.14 <2e-16 ***
## LastAuthorFemale1 0.05994 0.05210 1.15 0.25
## Year1997 0.22820 0.23955 0.95 0.34
## Year1998 0.20003 0.18815 1.06 0.29
## Year1999 -0.03724 0.14683 -0.25 0.80
## Year2000 0.04491 0.13286 0.34 0.74
## Year2001 0.03363 0.18632 0.18 0.86
## Year2002 0.18456 0.12155 1.52 0.13
## Year2003 0.11273 0.14871 0.76 0.45
## Year2004 0.19205 0.12138 1.58 0.11
## Year2005 -0.00113 0.12451 -0.01 0.99
## Year2006 0.01038 0.12106 0.09 0.93
```

```

## Year2007          -0.06212      0.12195    -0.51      0.61
## Year2008           0.08007      0.15385      0.52      0.60
## Year2009          -0.05346      0.14912    -0.36      0.72
## Year2010           0.02691      0.12987      0.21      0.84
## Year2011           0.02254      0.13220      0.17      0.86
## Year2012          -0.09577      0.14366    -0.67      0.51
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.054, Adjusted R-squared:  -0.00264
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 277 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.209  0.884  0.953  0.905  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      3.31e-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: 302"
## [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
##   19   20   32   29   25   30   41   19   27   24   38   30   27   35   22
## 2011 2012
##   26   35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    9    9   13   11   12   23   11   16   18   27   18   10   26   16
## 2011 2012

```

```
## 16 22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 8 9 12 10 12 22 10 15 18 25 17 9 24 16
## 2011 2012
## 15 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 = 13, df = 16, p-value = 0.7
```



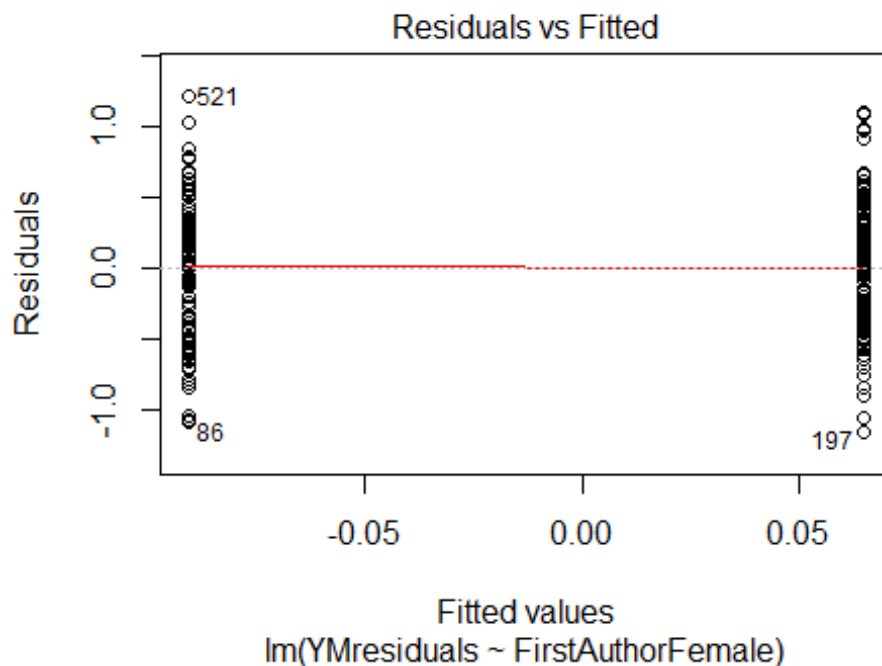
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.31, df = 1, p-value = 0.6

## [1] "Female first author team size 2018 geometric mean: 3.37628579092851"
## [1] "Male first author team size 2018 geometric mean: 3.16227766016838"

## 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 = 20, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.11214924268908"
## [1] "Male last author team size 2018 geometric mean: 4.04954688551703"

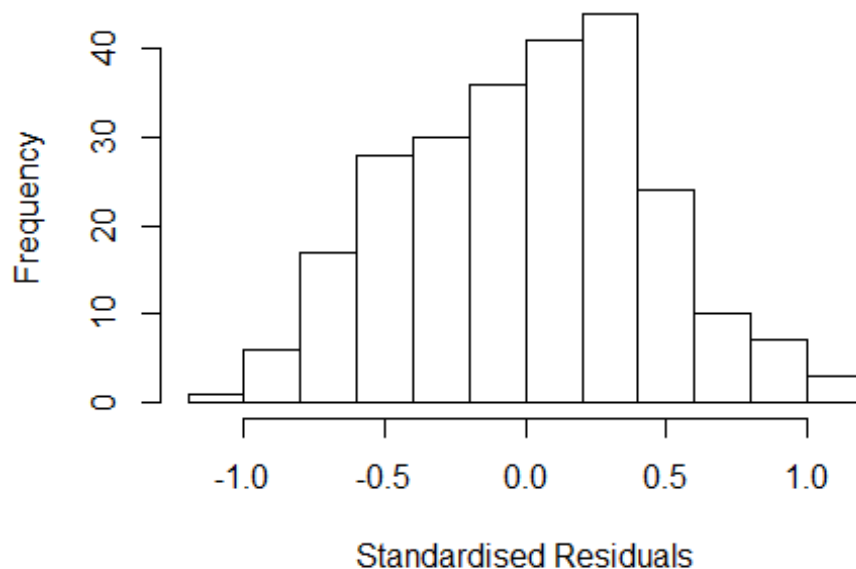
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 32, p-value = 0.3
## 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	1.528	1	1.236
LastAuthorFemale	1.401	1	1.184
UniqueAuthors	3.986	4	1.189
Year	6.058	16	1.058

Residuals from first and last author and team size



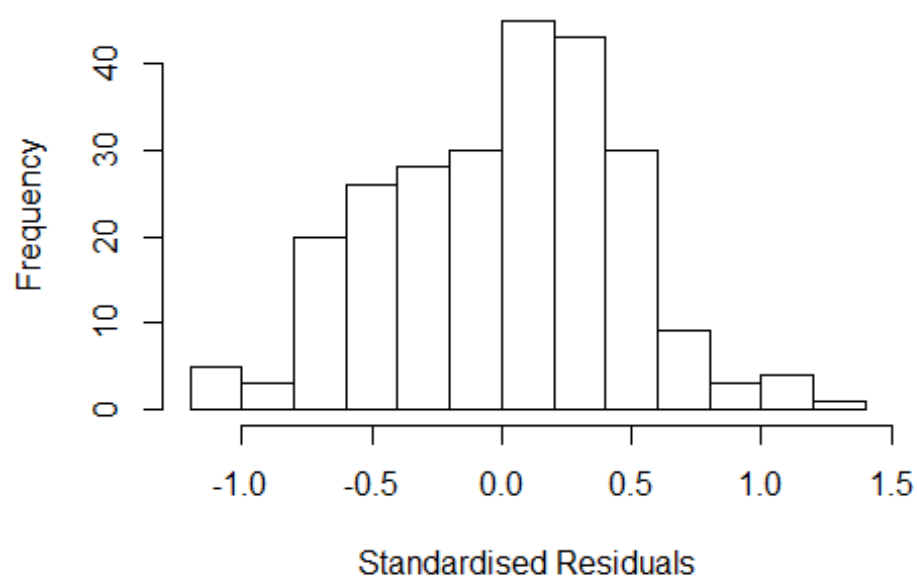
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.165 -0.325 0.026 0.290 1.174
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.87118 0.18422 4.73 4e-06 ***
## FirstAuthorFemale1 0.07786 0.06887 1.13 0.25947
## LastAuthorFemale1 -0.04553 0.06844 -0.67 0.50654
## UniqueAuthors2 0.29192 0.13238 2.21 0.02846 *
## UniqueAuthors3 0.29950 0.13552 2.21 0.02812 *
## UniqueAuthors4 0.38778 0.14945 2.59 0.01009 *
## UniqueAuthors5 0.48699 0.12785 3.81 0.00018 ***
## Year1997 0.24250 0.24581 0.99 0.32494
## Year1998 0.14534 0.24055 0.60 0.54632
## Year1999 0.00154 0.25417 0.01 0.99517
```

```

## Year2000          0.08783      0.19738      0.44  0.65676
## Year2001          0.24886      0.19627      1.27  0.20613
## Year2002         -0.12297      0.18812     -0.65  0.51399
## Year2003         -0.24000      0.18933     -1.27  0.20625
## Year2004         -0.42729      0.23155     -1.85  0.06631 .
## Year2005         -0.32038      0.18822     -1.70  0.09012 .
## Year2006         -0.07138      0.17332     -0.41  0.68084
## Year2007         -0.03599      0.18813     -0.19  0.84845
## Year2008         -0.06253      0.24775     -0.25  0.80097
## Year2009         -0.23645      0.18566     -1.27  0.20413
## Year2010         -0.17059      0.21635     -0.79  0.43125
## Year2011          0.00423      0.19319      0.02  0.98254
## Year2012         -0.04556      0.19362     -0.24  0.81417
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.45
## Multiple R-squared:  0.206, Adjusted R-squared:  0.128
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 231 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.475  0.882  0.947  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      4.05e-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 1.416 1      1.190
## LastAuthorFemale  1.271 1      1.128
## Year              1.652 16      1.016

```

Residuals from first and last author



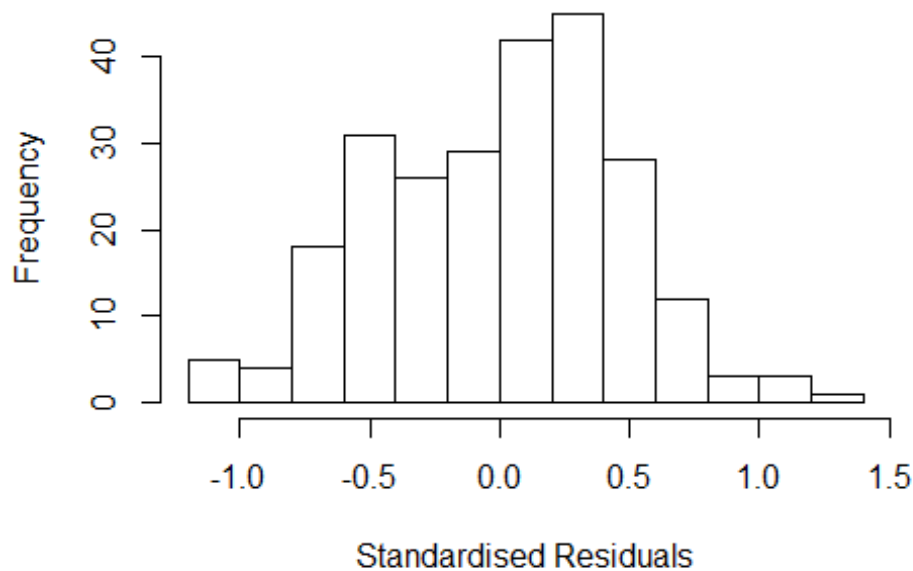
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1583 -0.3589 0.0364 0.3286 1.2209
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.05166 0.16896 6.22 2.3e-09 ***
## FirstAuthorFemale1 0.14745 0.07100 2.08 0.039 *
## LastAuthorFemale1 -0.03770 0.06877 -0.55 0.584
## Year1997 0.27738 0.26406 1.05 0.295
## Year1998 0.29290 0.26971 1.09 0.279
## Year1999 0.04878 0.25286 0.19 0.847
## Year2000 0.13605 0.21939 0.62 0.536
## Year2001 0.43078 0.19552 2.20 0.029 *
## Year2002 -0.01612 0.20440 -0.08 0.937
## Year2003 -0.10591 0.19248 -0.55 0.583
## Year2004 -0.36532 0.21979 -1.66 0.098 .
## Year2005 -0.17822 0.20713 -0.86 0.390
```

```

## Year2006      -0.00689    0.18713   -0.04    0.971
## Year2007      0.10660    0.18805    0.57    0.571
## Year2008      0.04938    0.26428    0.19    0.852
## Year2009     -0.12227    0.20238   -0.60    0.546
## Year2010     -0.08557    0.22873   -0.37    0.709
## Year2011      0.17550    0.19800    0.89    0.376
## Year2012      0.08753    0.20291    0.43    0.667
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.471
## Multiple R-squared:  0.142, Adjusted R-squared:  0.074
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 222 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.482  0.877  0.947  0.910  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      4.05e-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 1.382 1      1.176
## Year              1.382 16      1.010

```

Residuals from first author



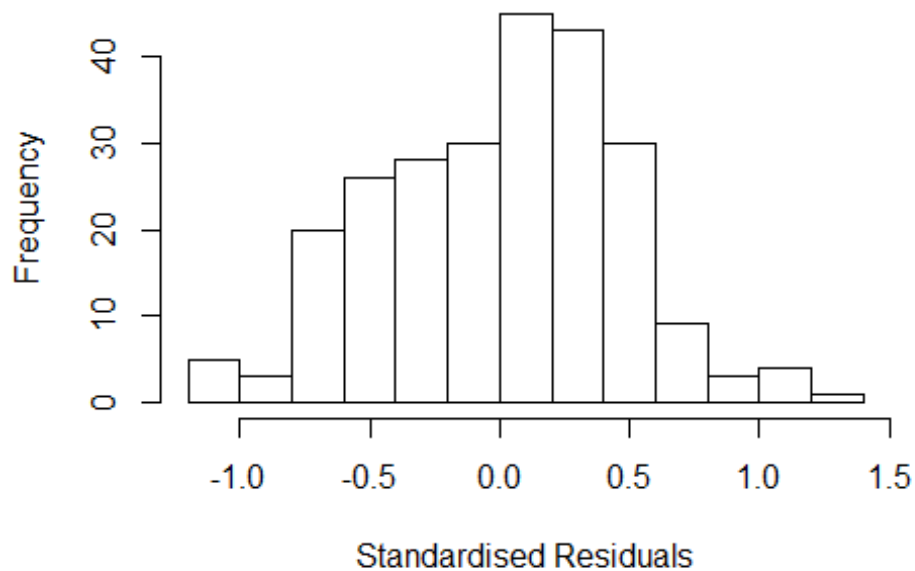
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1668 -0.3613 0.0347 0.3256 1.2363
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0536 0.1688 6.24 2.1e-09 ***
## FirstAuthorFemale1 0.1402 0.0700 2.00 0.046 *
## Year1997 0.2688 0.2610 1.03 0.304
## Year1998 0.2835 0.2662 1.06 0.288
## Year1999 0.0430 0.2532 0.17 0.865
## Year2000 0.1250 0.2171 0.58 0.565
## Year2001 0.4202 0.1944 2.16 0.032 *
## Year2002 -0.0271 0.2043 -0.13 0.895
## Year2003 -0.1181 0.1915 -0.62 0.538
## Year2004 -0.3727 0.2209 -1.69 0.093 .
## Year2005 -0.1962 0.2034 -0.96 0.336
## Year2006 -0.0123 0.1869 -0.07 0.948
```

```

## Year2007          0.0949      0.1861      0.51      0.611
## Year2008          0.0310      0.2589      0.12      0.905
## Year2009         -0.1387      0.2012     -0.69      0.491
## Year2010         -0.1029      0.2239     -0.46      0.646
## Year2011          0.1655      0.1968      0.84      0.401
## Year2012          0.0678      0.1983      0.34      0.733
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.141, Adjusted R-squared:  0.0771
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 223 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.468  0.882  0.942  0.909  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      4.05e-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 1.263 1      1.124
## Year            1.263 16      1.007

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2321 -0.3407 0.0169 0.3100 1.1884
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0868 0.1697 6.41 8.4e-10 ***
## LastAuthorFemale1 -0.0106 0.0681 -0.16 0.876
## Year1997 0.2325 0.2652 0.88 0.381
## Year1998 0.3255 0.2499 1.30 0.194
## Year1999 0.1041 0.2705 0.38 0.701
## Year2000 0.1757 0.2308 0.76 0.447
## Year2001 0.4734 0.1974 2.40 0.017 *
## Year2002 0.0055 0.2076 0.03 0.979
## Year2003 -0.0371 0.1935 -0.19 0.848
## Year2004 -0.3311 0.2244 -1.48 0.141
## Year2005 -0.1732 0.2163 -0.80 0.424
## Year2006 0.0393 0.1939 0.20 0.840
```

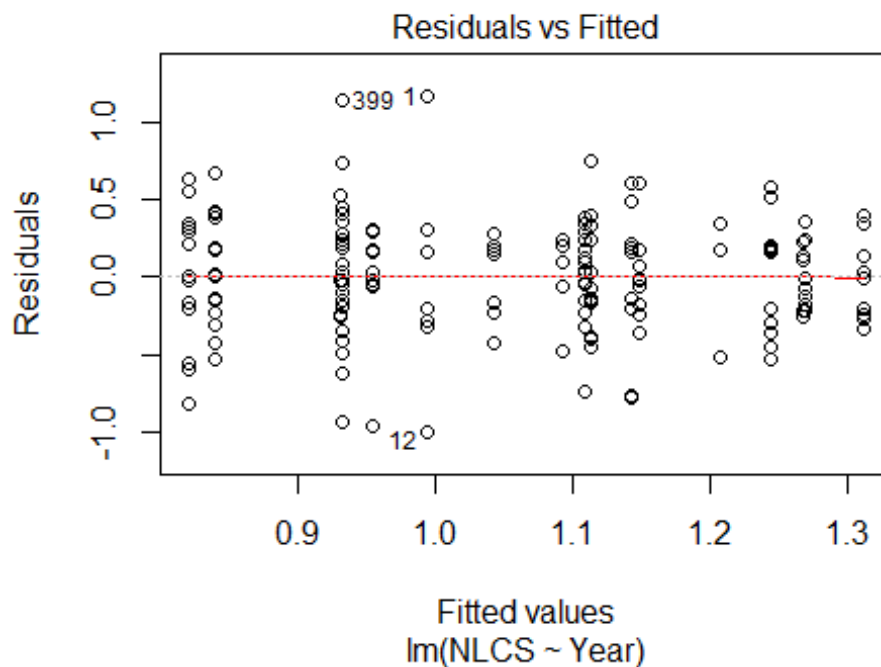
```

## Year2007          0.1454      0.1872      0.78      0.438
## Year2008          0.1305      0.2733      0.48      0.634
## Year2009         -0.0560      0.2011     -0.28      0.781
## Year2010         -0.0403      0.2172     -0.19      0.853
## Year2011          0.2305      0.2003      1.15      0.251
## Year2012          0.1438      0.2027      0.71      0.479
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.126, Adjusted R-squared:  0.0609
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 229 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.470  0.881  0.946  0.910  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      4.05e-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: 247"
## [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
##   12   26   18   24   10   20   18    8   13    9   15   15   20   18   20
## 2011 2012
##   25   35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8    9   11   16    5    5   10    3    7    5   13   10   12    7   11
## 2011 2012

```



```
## 17 24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 9 8 15 4 5 10 3 5 5 12 9 11 6 10
## 2011 2012
## 16 21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



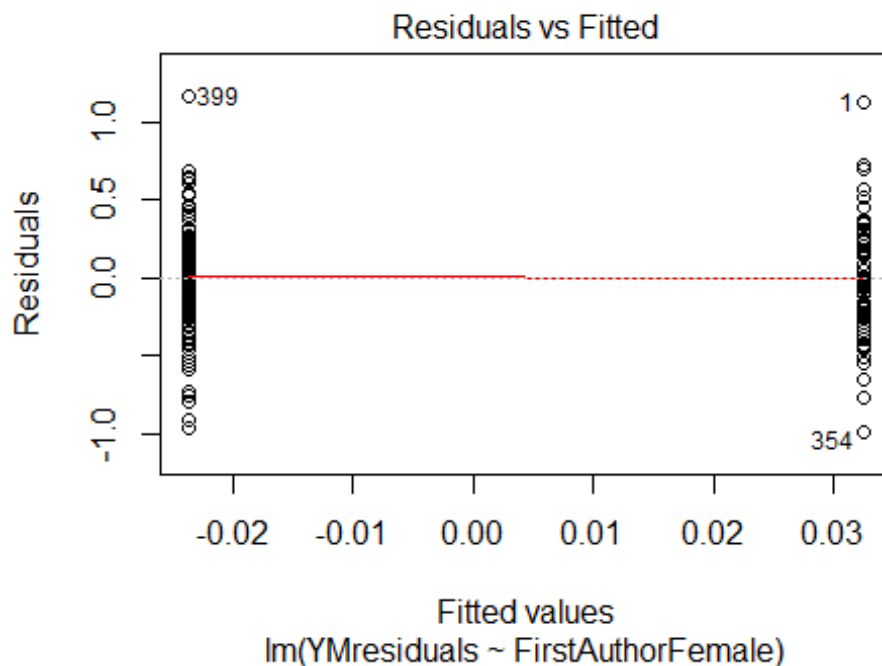
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.24, df = 1, p-value = 0.6

## [1] "Female first author team size 2018 geometric mean: 3.8328682607639"
## [1] "Male first author team size 2018 geometric mean: 4.3269826154684"

## 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 = 36, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.22417492183903"
## [1] "Male last author team size 2018 geometric mean: 3.9073156563647"

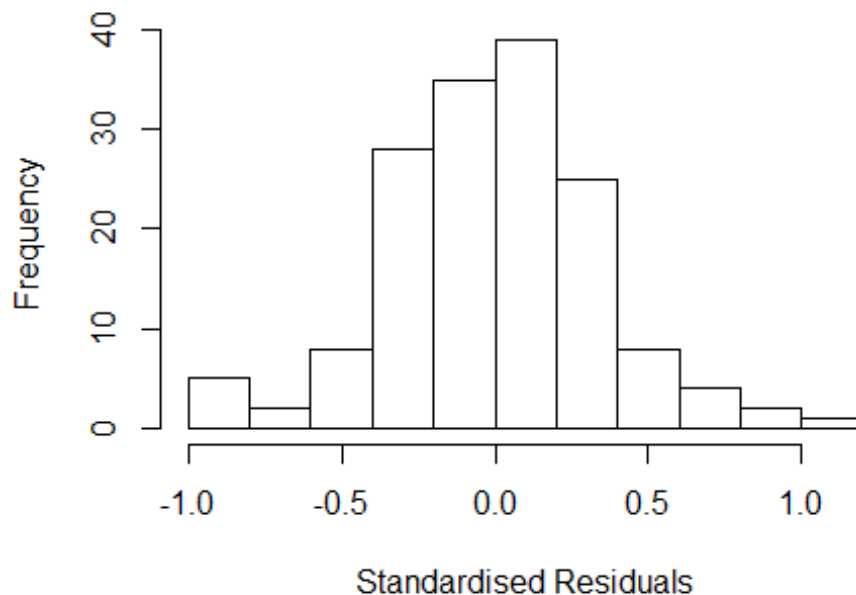
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 42, 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	2.307	1	1.519
LastAuthorFemale	1.856	1	1.362
UniqueAuthors	15.512	4	1.409
Year	40.066	16	1.122

Residuals from first and last author and team size



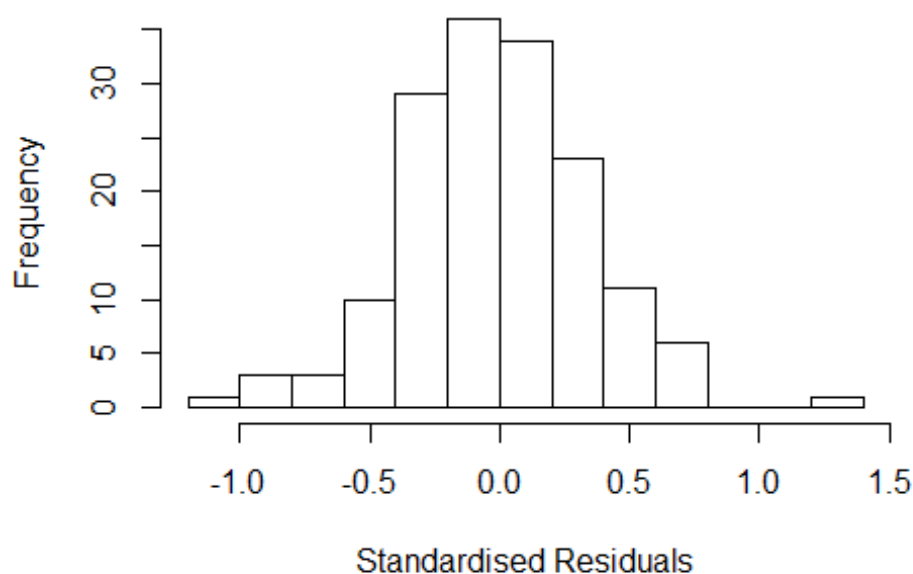
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9242 -0.2186  0.0013  0.2035  1.0941
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8077    0.1800   4.49 1.5e-05 ***
## FirstAuthorFemale1  0.0208    0.0679   0.31  0.76
## LastAuthorFemale1  0.0535    0.0669   0.80  0.42
## UniqueAuthors2   -0.0305    0.2006  -0.15  0.88
## UniqueAuthors3    0.1279    0.1816   0.70  0.48
## UniqueAuthors4    0.0923    0.1539   0.60  0.55
## UniqueAuthors5    0.2384    0.1735   1.37  0.17
## Year1997          0.2045    0.1930   1.06  0.29
## Year1998          0.2101    0.2469   0.85  0.40
## Year1999          0.1170    0.2026   0.58  0.56
```

```

## Year2000          0.1183      0.2788      0.42      0.67
## Year2001          0.3387      0.2083      1.63      0.11
## Year2002          0.2611      0.2562      1.02      0.31
## Year2003          0.1861      0.3311      0.56      0.57
## Year2004          0.1303      0.2098      0.62      0.54
## Year2005         -0.0579      0.2207     -0.26      0.79
## Year2006         -0.0905      0.2390     -0.38      0.71
## Year2007          0.3225      0.2001      1.61      0.11
## Year2008          0.0873      0.1908      0.46      0.65
## Year2009          0.2457      0.2092      1.17      0.24
## Year2010          0.0191      0.2048      0.09      0.93
## Year2011         -0.1339      0.2194     -0.61      0.54
## Year2012         -0.0502      0.1912     -0.26      0.79
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.325
## Multiple R-squared:  0.238, Adjusted R-squared:  0.113
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.233  0.891  0.952  0.898  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      6.37e-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 1.792 1      1.339
## LastAuthorFemale  1.528 1      1.236
## Year              2.414 16      1.028

```

Residuals from first and last author



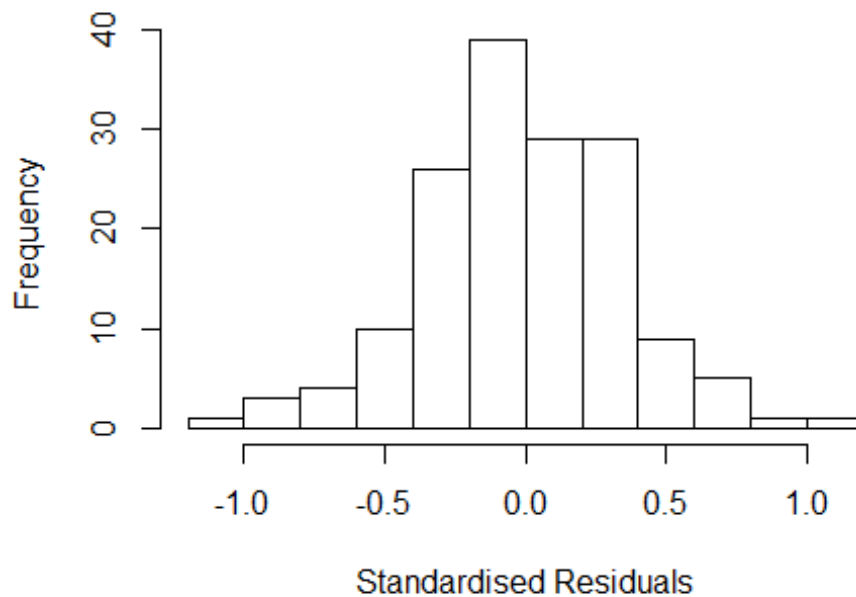
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0469 -0.2141 -0.0154  0.2127  1.2097
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9028    0.1881   4.80 4.1e-06 ***
## FirstAuthorFemale1 0.0485    0.0662   0.73  0.466
## LastAuthorFemale1 0.0465    0.0650   0.72  0.475
## Year1997        0.1923    0.2067   0.93  0.354
## Year1998        0.2921    0.2632   1.11  0.269
## Year1999        0.1796    0.2003   0.90  0.372
## Year2000        0.1725    0.2528   0.68  0.496
## Year2001        0.3472    0.2058   1.69  0.094 .
## Year2002        0.2645    0.2488   1.06  0.290
## Year2003        0.3088    0.3158   0.98  0.330
## Year2004        0.0845    0.2356   0.36  0.720
## Year2005       -0.0259    0.2333  -0.11  0.912
```

```

## Year2006          -0.0767      0.2468   -0.31    0.756
## Year2007           0.3937      0.2100    1.87    0.063 .
## Year2008           0.0945      0.2097    0.45    0.653
## Year2009           0.2559      0.1983    1.29    0.199
## Year2010           0.0490      0.2010    0.24    0.808
## Year2011          -0.1292      0.2095   -0.62    0.538
## Year2012           0.0101      0.2001    0.05    0.960
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.343
## Multiple R-squared:  0.171, Adjusted R-squared:  0.0626
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 145 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.189  0.883  0.962  0.907  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      6.37e-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 1.683 1      1.297
## Year              1.683 16      1.016

```

Residuals from first author



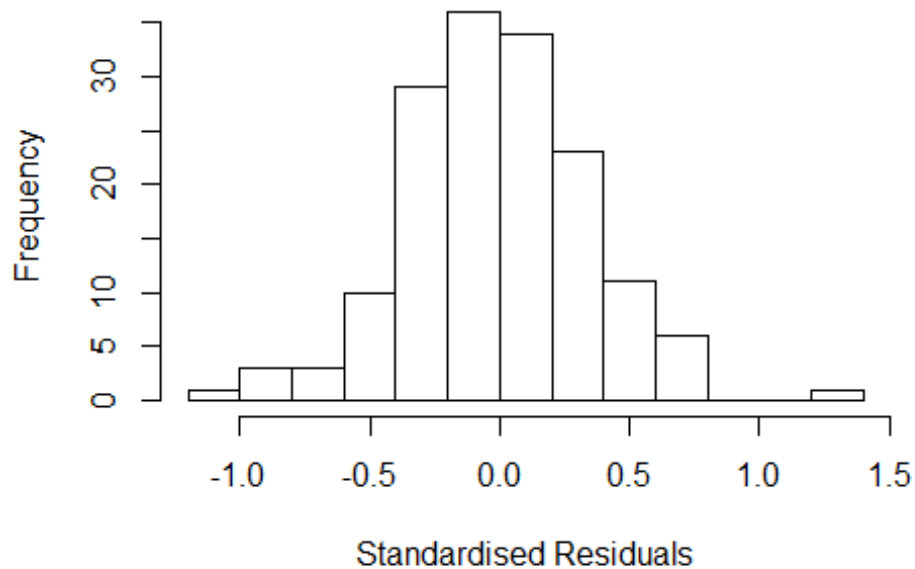
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0218 -0.2205 -0.0215  0.2224  1.1961
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9116    0.1925   4.74 5.3e-06 ***
## FirstAuthorFemale1 0.0533    0.0647   0.82  0.411
## Year1997        0.2099    0.2106   1.00  0.321
## Year1998        0.2956    0.2686   1.10  0.273
## Year1999        0.1755    0.2045   0.86  0.392
## Year2000        0.1633    0.2568   0.64  0.526
## Year2001        0.3465    0.2129   1.63  0.106
## Year2002        0.2586    0.2514   1.03  0.305
## Year2003        0.2978    0.3247   0.92  0.361
## Year2004        0.1024    0.2413   0.42  0.672
## Year2005       -0.0382    0.2355  -0.16  0.871
## Year2006       -0.0753    0.2494  -0.30  0.763
```

```

## Year2007          0.3978      0.2122      1.87      0.063 .
## Year2008          0.0965      0.2163      0.45      0.656
## Year2009          0.2597      0.2026      1.28      0.202
## Year2010          0.0569      0.2075      0.27      0.784
## Year2011         -0.1253      0.2159     -0.58      0.563
## Year2012          0.0197      0.2062      0.10      0.924
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.167, Adjusted R-squared:  0.0653
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 144 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.196  0.888  0.956  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      6.37e-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 1.468 1      1.212
## Year              1.468 16      1.012

```


Residuals from last author



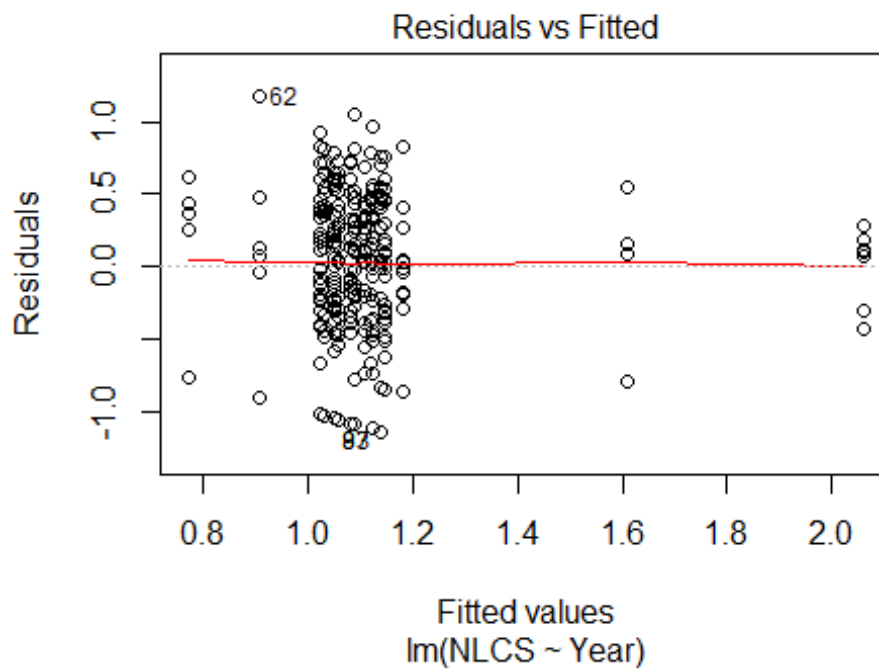
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0285 -0.2224 -0.0167 0.2092 1.2488
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.91222 0.18503 4.93 2.3e-06 ***
## LastAuthorFemale1 0.05244 0.06330 0.83 0.409
## Year1997 0.18985 0.20219 0.94 0.349
## Year1998 0.28077 0.26140 1.07 0.285
## Year1999 0.18061 0.19732 0.92 0.362
## Year2000 0.17076 0.24289 0.70 0.483
## Year2001 0.34691 0.20415 1.70 0.092 .
## Year2002 0.28653 0.24577 1.17 0.246
## Year2003 0.32683 0.31070 1.05 0.295
## Year2004 0.10901 0.22890 0.48 0.635
## Year2005 -0.00187 0.22305 -0.01 0.993
## Year2006 -0.06584 0.24947 -0.26 0.792
```

```

## Year2007      0.40421      0.20649      1.96      0.052 .
## Year2008      0.11682      0.20637      0.57      0.572
## Year2009      0.27606      0.19352      1.43      0.156
## Year2010      0.06388      0.19668      0.32      0.746
## Year2011     -0.10444      0.20145     -0.52      0.605
## Year2012      0.01501      0.19700      0.08      0.939
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.34
## Multiple R-squared:  0.168, Adjusted R-squared:  0.0666
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 146 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.150  0.887   0.956   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      6.37e-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: 157"
## [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
##   10   15   9   12   11   21   18   22   38   27   27   26   36   27   24
## 2011 2012
##   26   28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    7    8    9    7   16   12   16   32   20   23   21   24   19   20
## 2011 2012

```

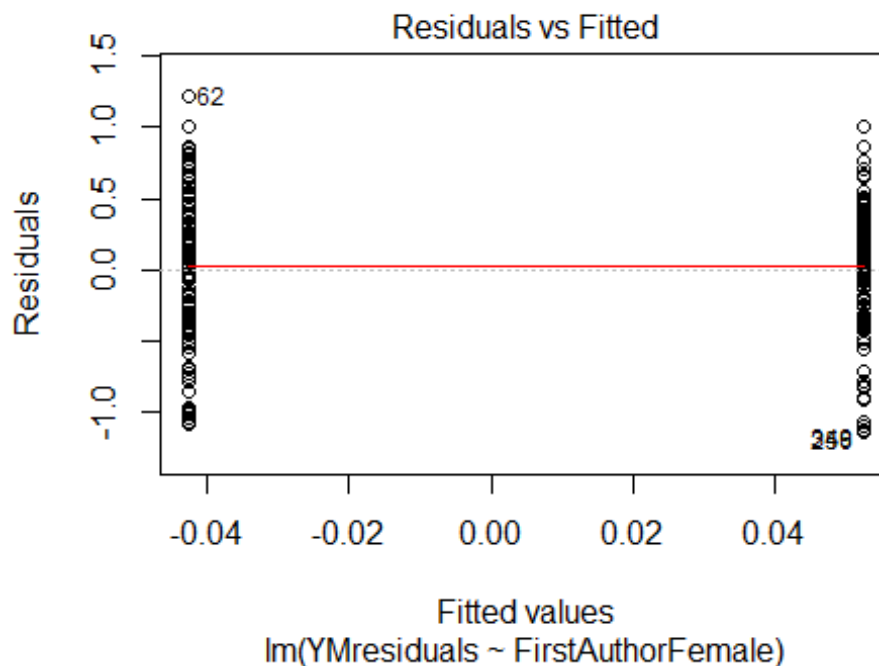
```
## 15 23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 7 8 9 7 16 11 15 29 19 17 21 23 18 19
## 2011 2012
## 15 23
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 19, df = 16, p-value = 0.3
```



```
##
## 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: 4.85961686300321"
## [1] "Male first author team size 2018 geometric mean: 2.80294088872968"
## 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 = 97, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.09889961704886"
## [1] "Male last author team size 2018 geometric mean: 2.99255573947769"

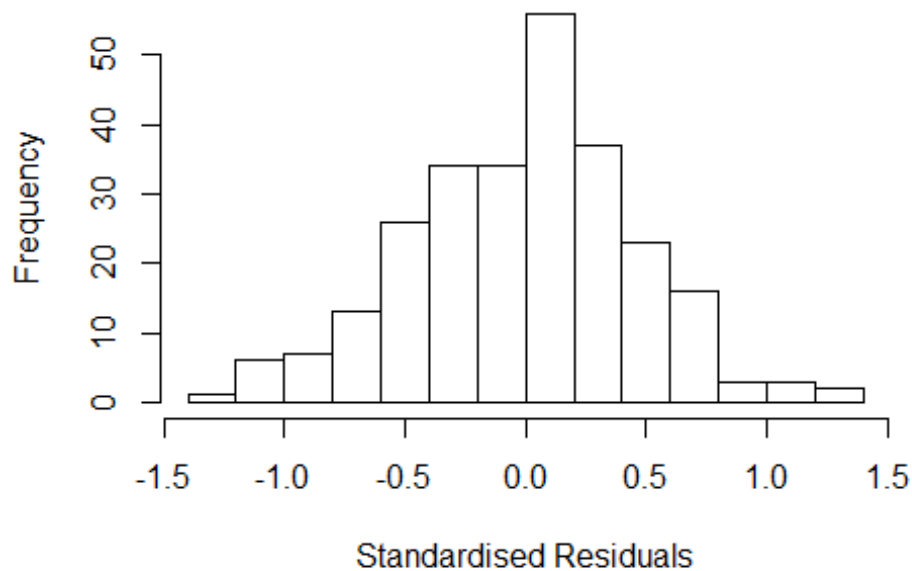
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 85, 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	1.498	1	1.224
LastAuthorFemale	1.509	1	1.228
UniqueAuthors	3.515	4	1.170
Year	4.863	16	1.051

Residuals from first and last author and team size



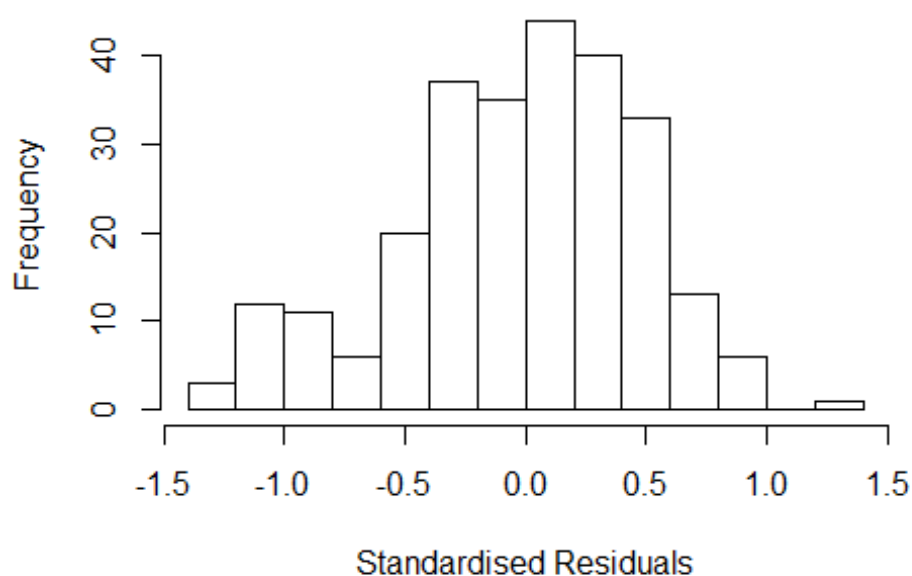
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3388 -0.3256  0.0263  0.3027  1.2242
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.34628    0.24082     5.59 6.2e-08 ***
## FirstAuthorFemale1  0.00234    0.06681     0.04  0.9721
## LastAuthorFemale1 -0.09763    0.07682    -1.27  0.2050
## UniqueAuthors2     0.33833    0.16692     2.03  0.0438 *
## UniqueAuthors3     0.48602    0.15437     3.15  0.0019 **
## UniqueAuthors4     0.48184    0.15150     3.18  0.0017 **
## UniqueAuthors5     0.60028    0.14848     4.04 7.1e-05 ***
## Year1997           0.25742    0.22504     1.14  0.2538
## Year1998          -0.90059    0.30119    -2.99  0.0031 **
## Year1999          -0.57945    0.24590    -2.36  0.0193 *
```

```

## Year2000      -0.82476    0.28457   -2.90    0.0041 **
## Year2001      -0.59039    0.26005   -2.27    0.0241 *
## Year2002      -0.60269    0.20952   -2.88    0.0044 **
## Year2003      -0.66977    0.22436   -2.99    0.0031 **
## Year2004      -0.64127    0.21304   -3.01    0.0029 **
## Year2005      -0.61009    0.27013   -2.26    0.0248 *
## Year2006      -0.69526    0.23125   -3.01    0.0029 **
## Year2007      -0.65665    0.24929   -2.63    0.0090 **
## Year2008      -0.63604    0.21910   -2.90    0.0040 **
## Year2009      -0.67733    0.22702   -2.98    0.0031 **
## Year2010      -0.60785    0.21266   -2.86    0.0046 **
## Year2011      -0.76692    0.24148   -3.18    0.0017 **
## Year2012      -0.65982    0.21687   -3.04    0.0026 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.232, Adjusted R-squared:  0.161
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 233 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.375  0.867  0.949  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          3.83e-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 1.248 1          1.117
## LastAuthorFemale  1.798 1          1.341
## Year              2.090 16          1.023

```

Residuals from first and last author



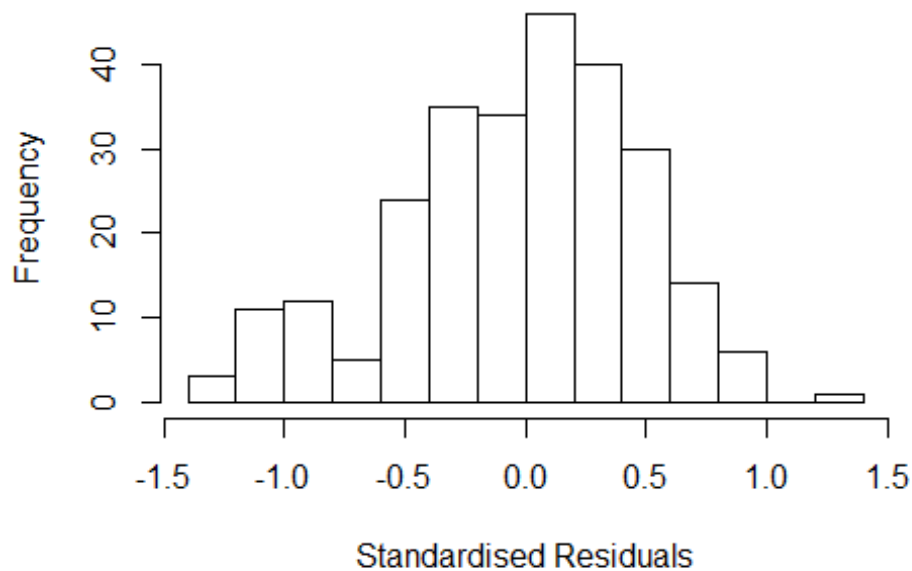
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2238 -0.3106 0.0362 0.3311 1.2176
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6557 0.2584 6.41 7.6e-10 ***
## FirstAuthorFemale1 0.0918 0.0667 1.38 0.170
## LastAuthorFemale1 -0.0716 0.0936 -0.76 0.445
## Year1997 0.3966 0.2752 1.44 0.151
## Year1998 -0.8575 0.3896 -2.20 0.029 *
## Year1999 -0.5694 0.3225 -1.77 0.079 .
## Year2000 -0.7893 0.4124 -1.91 0.057 .
## Year2001 -0.4320 0.3092 -1.40 0.164
## Year2002 -0.5081 0.2867 -1.77 0.078 .
## Year2003 -0.5830 0.2780 -2.10 0.037 *
## Year2004 -0.5936 0.2663 -2.23 0.027 *
## Year2005 -0.5328 0.2944 -1.81 0.072 .
```

```

## Year2006          -0.5768      0.2975    -1.94     0.054 .
## Year2007          -0.5587      0.2857    -1.96     0.052 .
## Year2008          -0.5540      0.2731    -2.03     0.044 *
## Year2009          -0.5907      0.2780    -2.12     0.035 *
## Year2010          -0.5092      0.2685    -1.90     0.059 .
## Year2011          -0.6458      0.2926    -2.21     0.028 *
## Year2012          -0.5279      0.2744    -1.92     0.056 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.457
## Multiple R-squared:  0.138, Adjusted R-squared:  0.0737
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 240 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.453  0.862  0.946  0.893  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      3.83e-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 1.217 1      1.103
## Year              1.217 16      1.006

```


Residuals from first author



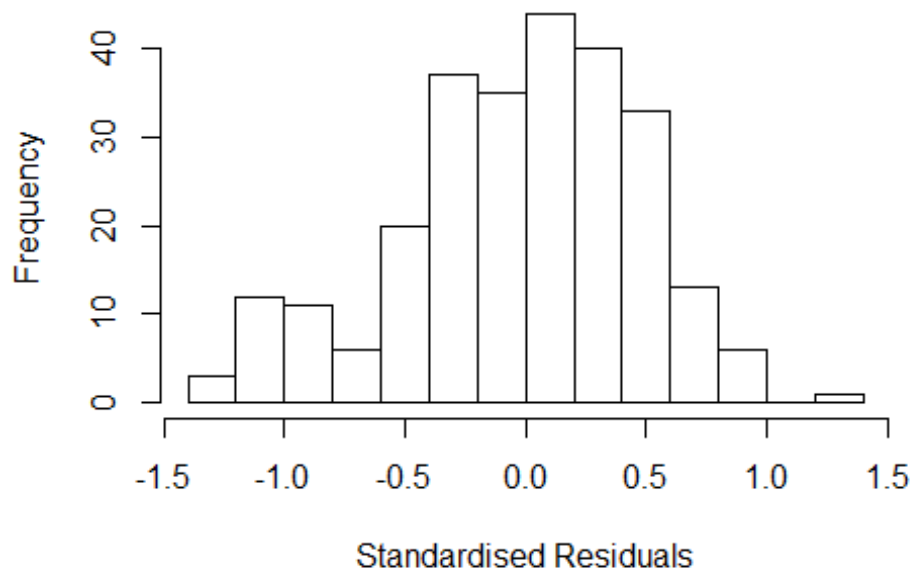
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2172 -0.3135 0.0285 0.3412 1.2184
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6316 0.2493 6.54 3.5e-10 ***
## FirstAuthorFemale1 0.0944 0.0666 1.42 0.157
## Year1997 0.4091 0.2690 1.52 0.130
## Year1998 -0.8401 0.3980 -2.11 0.036 *
## Year1999 -0.5670 0.3229 -1.76 0.080 .
## Year2000 -0.7660 0.4076 -1.88 0.061 .
## Year2001 -0.4144 0.3023 -1.37 0.172
## Year2002 -0.4862 0.2791 -1.74 0.083 .
## Year2003 -0.5599 0.2705 -2.07 0.040 *
## Year2004 -0.5850 0.2609 -2.24 0.026 *
## Year2005 -0.5124 0.2860 -1.79 0.074 .
## Year2006 -0.5516 0.2886 -1.91 0.057 .
```

```

## Year2007          -0.5562      0.2826    -1.97      0.050 .
## Year2008          -0.5395      0.2680    -2.01      0.045 *
## Year2009          -0.5776      0.2724    -2.12      0.035 *
## Year2010          -0.5155      0.2648    -1.95      0.053 .
## Year2011          -0.6423      0.2912    -2.21      0.028 *
## Year2012          -0.5330      0.2717    -1.96      0.051 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.457
## Multiple R-squared:  0.135, Adjusted R-squared:  0.0749
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 245 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.456  0.867   0.945   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      3.83e-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 1.635 1          1.279
## Year            1.635 16          1.015

```

Residuals from last author



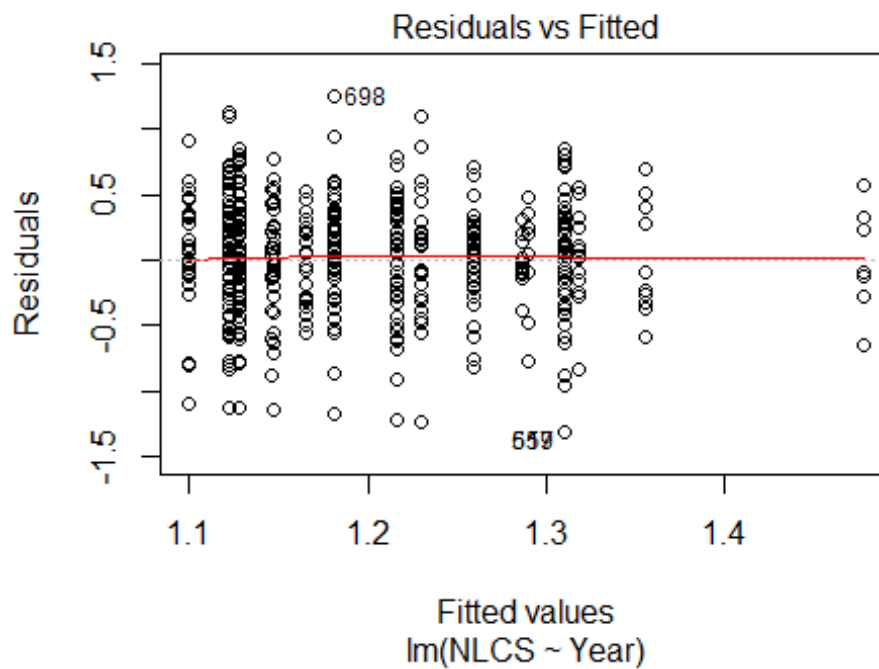
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2413 -0.3279 0.0388 0.3461 1.1901
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6746 0.2740 6.11 3.9e-09 ***
## LastAuthorFemale1 -0.0767 0.0902 -0.85 0.396
## Year1997 0.4024 0.2877 1.40 0.163
## Year1998 -0.8388 0.3826 -2.19 0.029 *
## Year1999 -0.5381 0.3229 -1.67 0.097 .
## Year2000 -0.7807 0.4172 -1.87 0.062 .
## Year2001 -0.4333 0.3253 -1.33 0.184
## Year2002 -0.4836 0.2983 -1.62 0.106
## Year2003 -0.5665 0.2953 -1.92 0.056 .
## Year2004 -0.5757 0.2819 -2.04 0.042 *
## Year2005 -0.5102 0.3111 -1.64 0.102
## Year2006 -0.5697 0.3134 -1.82 0.070 .
```

```

## Year2007          -0.5106      0.2991    -1.71     0.089 .
## Year2008          -0.5011      0.2880    -1.74     0.083 .
## Year2009          -0.5638      0.2928    -1.93     0.055 .
## Year2010          -0.4887      0.2832    -1.73     0.086 .
## Year2011          -0.6237      0.3014    -2.07     0.040 *
## Year2012          -0.5089      0.2882    -1.77     0.079 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.482
## Multiple R-squared:  0.128, Adjusted R-squared:  0.0667
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 234 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.487  0.873   0.946   0.900   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      3.83e-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: 261"
## [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
##   23   16   15   22   23   32   18   29   44   47   41   46   50   70   75
## 2011 2012
##   98   86
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7   10    8   11   12   14   12   22   30   28   30   36   34   48   56
## 2011 2012

```

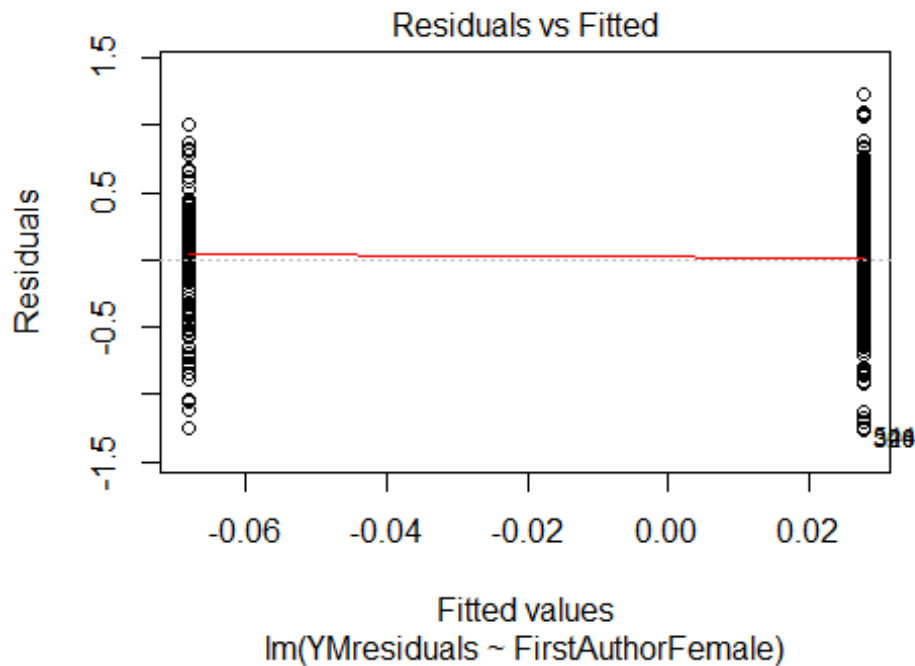
```
## 71 67
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 10 8 9 11 13 10 17 29 22 26 34 32 42 49
## 2011 2012
## 62 57
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 25, df = 16, p-value = 0.08
```



```
##
## 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: 4.31847854958547"
## [1] "Male first author team size 2018 geometric mean: 3.71446621413842"
## 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 = 320, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.51276529679714"
## [1] "Male last author team size 2018 geometric mean: 4.00648452170791"

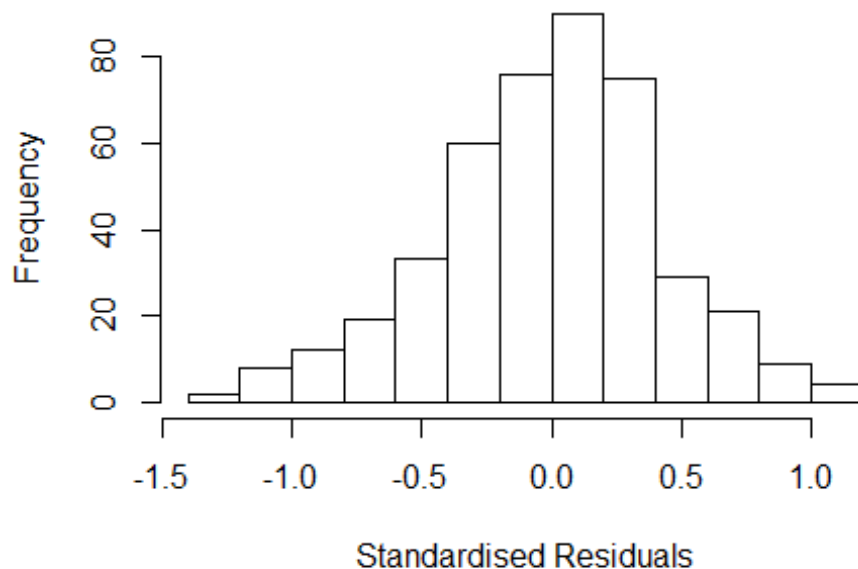
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 120, 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	1.358	1	1.165
LastAuthorFemale	1.209	1	1.100
UniqueAuthors	2.216	4	1.105
Year	3.016	16	1.035

Residuals from first and last author and team size



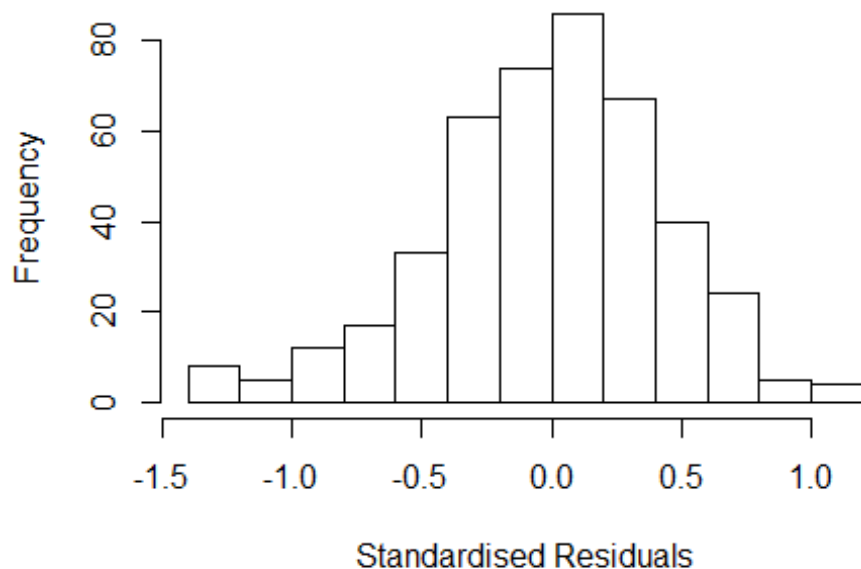
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3547 -0.2851  0.0137  0.2636  1.0793
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4097     0.1861    7.57 2.4e-13 ***
## FirstAuthorFemale1 -0.0416     0.0494   -0.84 0.39995
## LastAuthorFemale1 -0.0155     0.0511   -0.30 0.76157
## UniqueAuthors2      0.1822     0.0726    2.51 0.01245 *
## UniqueAuthors3      0.2825     0.0734    3.85 0.00014 ***
## UniqueAuthors4      0.2489     0.0768    3.24 0.00128 **
## UniqueAuthors5      0.3608     0.0850    4.24 2.7e-05 ***
## Year1997          -0.2437     0.2143   -1.14 0.25612
## Year1998          -0.2017     0.2688   -0.75 0.45357
## Year1999          -0.3062     0.2021   -1.51 0.13058
```

```

## Year2000          -0.4153      0.2166    -1.92   0.05586 .
## Year2001          -0.2107      0.2041    -1.03   0.30247
## Year2002          -0.5306      0.2420    -2.19   0.02890 *
## Year2003          -0.4593      0.2132    -2.15   0.03177 *
## Year2004          -0.3954      0.2052    -1.93   0.05467 .
## Year2005          -0.4110      0.2276    -1.81   0.07167 .
## Year2006          -0.4970      0.2133    -2.33   0.02032 *
## Year2007          -0.3205      0.1989    -1.61   0.10775
## Year2008          -0.3250      0.2056    -1.58   0.11478
## Year2009          -0.2874      0.2068    -1.39   0.16539
## Year2010          -0.4158      0.2038    -2.04   0.04198 *
## Year2011          -0.4922      0.1994    -2.47   0.01396 *
## Year2012          -0.4682      0.2025    -2.31   0.02123 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.41
## Multiple R-squared:  0.101, Adjusted R-squared:  0.0532
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 37 weights are ~= 1. The remaining 401 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.252  0.868  0.949   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          2.28e-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 1.234 1          1.111
## LastAuthorFemale  1.159 1          1.077
## Year              1.316 16          1.009

```


Residuals from first and last author



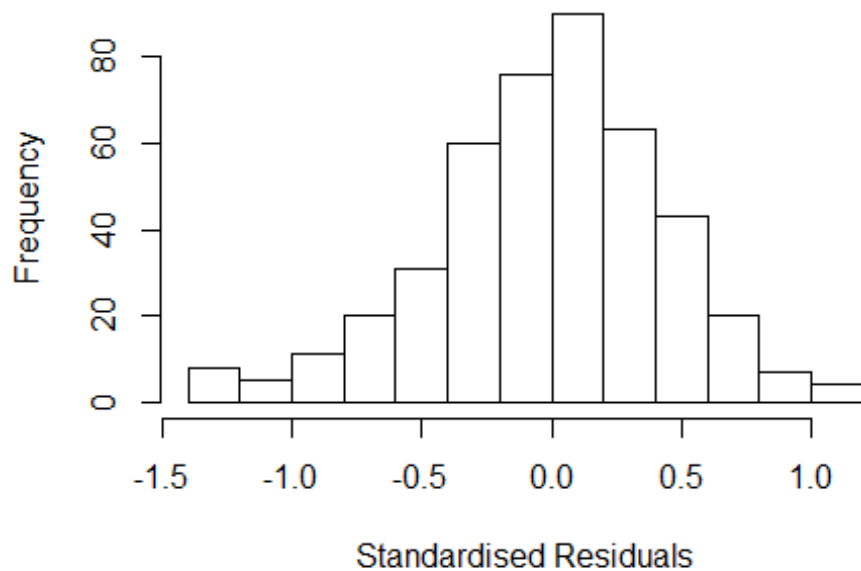
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3056 -0.2848  0.0242  0.2807  1.1893
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4957     0.1607   9.31  <2e-16 ***
## FirstAuthorFemale1 -0.0428     0.0496  -0.86   0.388
## LastAuthorFemale1 -0.0386     0.0519  -0.74   0.458
## Year1997          -0.1547     0.2193  -0.71   0.481
## Year1998          -0.1497     0.2215  -0.68   0.499
## Year1999          -0.2087     0.1717  -1.22   0.225
## Year2000          -0.2952     0.1905  -1.55   0.122
## Year2001          -0.1057     0.1792  -0.59   0.555
## Year2002          -0.3804     0.2192  -1.74   0.083 .
## Year2003          -0.3607     0.1855  -1.94   0.053 .
## Year2004          -0.2736     0.1766  -1.55   0.122
## Year2005          -0.2754     0.2034  -1.35   0.176
```

```

## Year2006          -0.3463      0.1801    -1.92     0.055 .
## Year2007          -0.1953      0.1725    -1.13     0.258
## Year2008          -0.1901      0.1782    -1.07     0.287
## Year2009          -0.1234      0.1772    -0.70     0.487
## Year2010          -0.2510      0.1743    -1.44     0.151
## Year2011          -0.3568      0.1717    -2.08     0.038 *
## Year2012          -0.3142      0.1738    -1.81     0.071 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.0461, Adjusted R-squared:  0.00514
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 400 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.312  0.860   0.951   0.897   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      2.28e-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 1.168 1      1.081
## Year              1.168 16      1.005

```

Residuals from first author



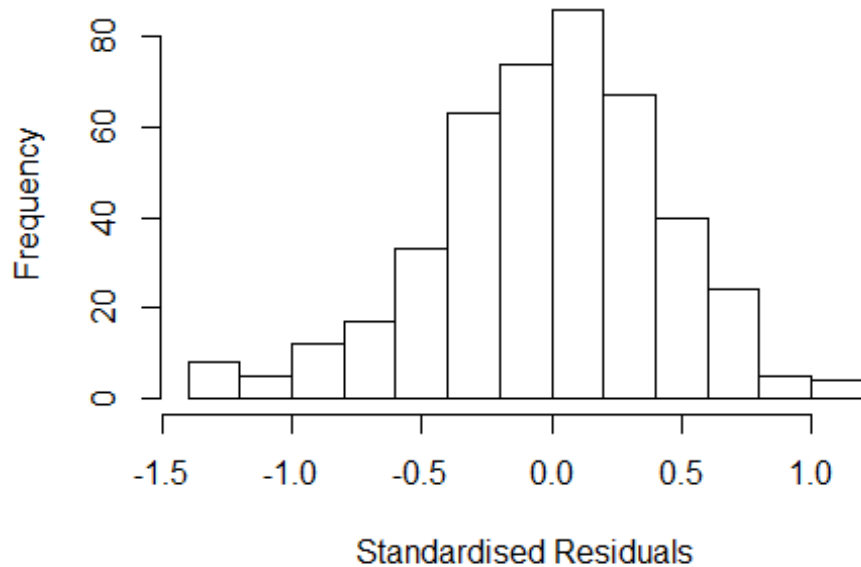
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3111 -0.2937 0.0271 0.2827 1.1990
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4918 0.1589 9.39 <2e-16 ***
## FirstAuthorFemale1 -0.0570 0.0493 -1.16 0.249
## Year1997 -0.1509 0.2181 -0.69 0.489
## Year1998 -0.1534 0.2199 -0.70 0.486
## Year1999 -0.2077 0.1703 -1.22 0.223
## Year2000 -0.2984 0.1883 -1.58 0.114
## Year2001 -0.1051 0.1769 -0.59 0.553
## Year2002 -0.3749 0.2182 -1.72 0.086 .
## Year2003 -0.3574 0.1846 -1.94 0.054 .
## Year2004 -0.2760 0.1753 -1.57 0.116
## Year2005 -0.2700 0.2019 -1.34 0.182
## Year2006 -0.3519 0.1784 -1.97 0.049 *
```

```

## Year2007          -0.1961      0.1712    -1.15     0.253
## Year2008          -0.1926      0.1768    -1.09     0.277
## Year2009          -0.1238      0.1755    -0.71     0.481
## Year2010          -0.2568      0.1724    -1.49     0.137
## Year2011          -0.3583      0.1700    -2.11     0.036 *
## Year2012          -0.3181      0.1720    -1.85     0.065 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.418
## Multiple R-squared:  0.045, Adjusted R-squared:  0.00629
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 406 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.305  0.863   0.951   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      2.28e-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 1.092 1          1.045
## Year            1.092 16          1.003

```

Residuals from last author



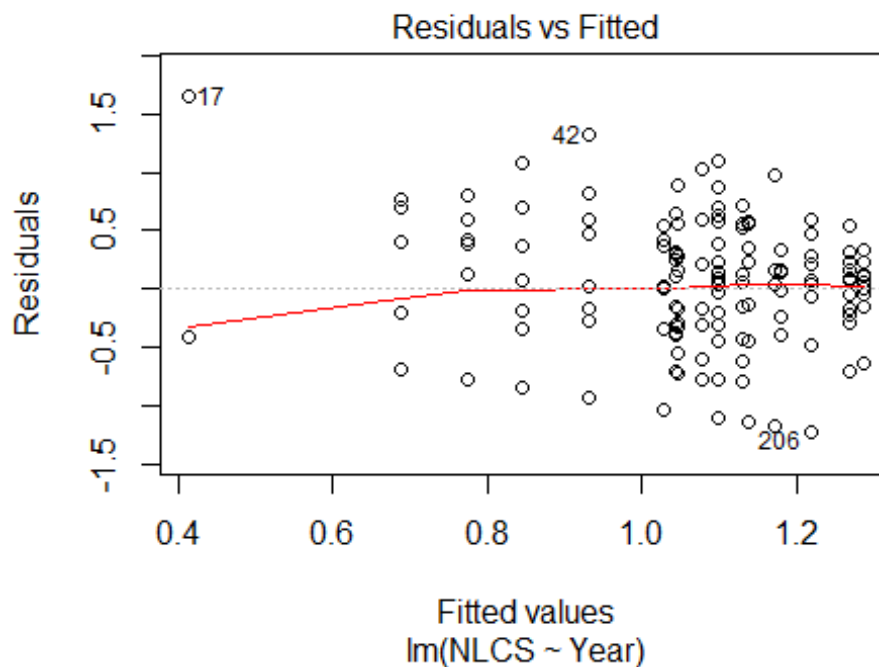
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3096 -0.2923 0.0136 0.2715 1.1950
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4915 0.1588 9.39 <2e-16 ***
## LastAuthorFemale1 -0.0551 0.0515 -1.07 0.285
## Year1997 -0.1506 0.2181 -0.69 0.490
## Year1998 -0.1536 0.2199 -0.70 0.485
## Year1999 -0.2076 0.1700 -1.22 0.223
## Year2000 -0.2881 0.1892 -1.52 0.129
## Year2001 -0.1067 0.1777 -0.60 0.549
## Year2002 -0.3816 0.2178 -1.75 0.081 .
## Year2003 -0.3609 0.1833 -1.97 0.050 *
## Year2004 -0.2793 0.1748 -1.60 0.111
## Year2005 -0.2852 0.2003 -1.42 0.155
## Year2006 -0.3440 0.1785 -1.93 0.055 .
```

```

## Year2007          -0.2042      0.1701    -1.20      0.231
## Year2008          -0.1913      0.1762    -1.09      0.278
## Year2009          -0.1267      0.1754    -0.72      0.470
## Year2010          -0.2525      0.1729    -1.46      0.145
## Year2011          -0.3659      0.1697    -2.16      0.032 *
## Year2012          -0.3169      0.1723    -1.84      0.067 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.418
## Multiple R-squared:  0.0447, Adjusted R-squared:  0.00605
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 406 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.305  0.863  0.952  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      2.28e-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: 438"
## [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
##   11   7  15  15  15  19  17  12  18  11  11  13  23  17  20
## 2011 2012
##   26  19
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8    5    4   10    7    7    7    8    9    8    7   10   12   11    9
## 2011 2012

```

```
## 20 11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 5 3 9 6 7 6 8 8 7 6 10 10 11 6
## 2011 2012
## 19 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 25, df = 16, p-value = 0.08
```



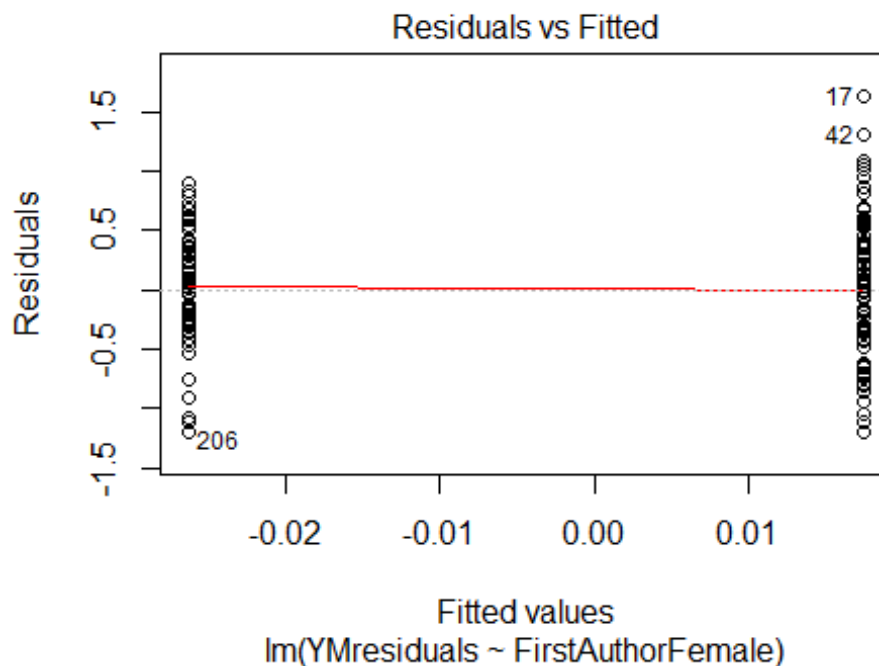
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.2, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 3.23124684238972"
## [1] "Male first author team size 2018 geometric mean: 4.80095151939279"

## 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 = 22, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.23124684238972"
## [1] "Male last author team size 2018 geometric mean: 4.80095151939279"

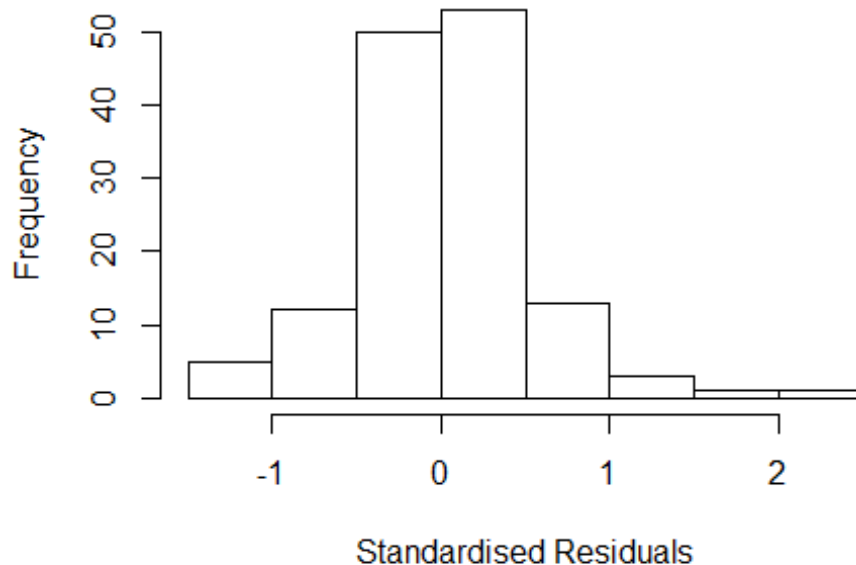
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 22, 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	2.711	1	1.646
LastAuthorFemale	2.091	1	1.446
UniqueAuthors	41.358	4	1.592
Year	97.901	16	1.154

Residuals from first and last author and team size



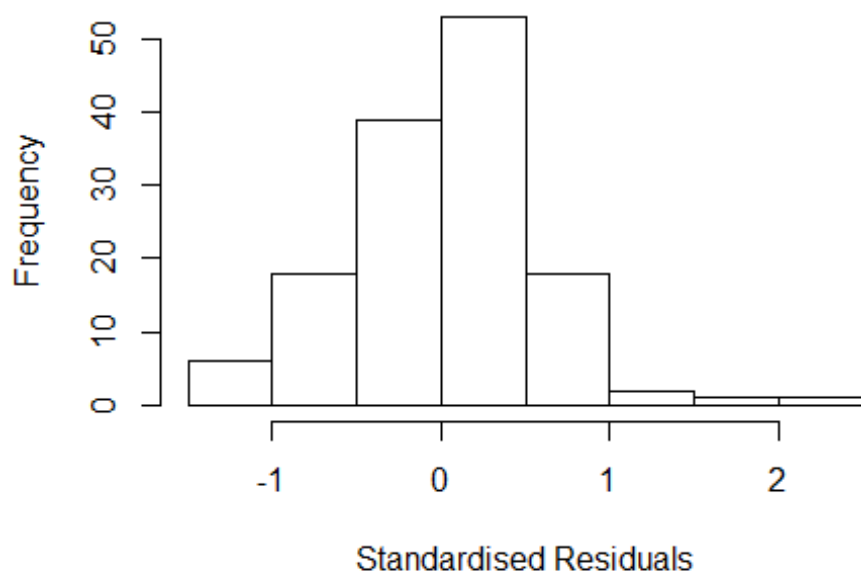
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.32566 -0.33576 0.00626 0.28854 2.23937
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3582 0.3474 1.03 0.305
## FirstAuthorFemale1 -0.1316 0.1149 -1.15 0.254
## LastAuthorFemale1 0.1966 0.0971 2.03 0.045 *
## UniqueAuthors2 0.5425 0.2621 2.07 0.041 *
## UniqueAuthors3 0.1817 0.2428 0.75 0.456
## UniqueAuthors4 0.4967 0.2393 2.08 0.040 *
## UniqueAuthors5 0.5647 0.2606 2.17 0.032 *
## Year1997 -0.5355 0.3147 -1.70 0.091 .
## Year1998 0.5771 1.3590 0.42 0.672
## Year1999 0.0535 0.3790 0.14 0.888
```

```

## Year2000          0.4053      0.3282      1.24      0.219
## Year2001          0.3709      0.3632      1.02      0.309
## Year2002          0.4733      0.3104      1.52      0.130
## Year2003         -0.0811      0.3719     -0.22      0.828
## Year2004          0.6031      0.3005      2.01      0.047 *
## Year2005          0.4028      0.3787      1.06      0.290
## Year2006          0.4085      0.3637      1.12      0.264
## Year2007          0.2764      0.3012      0.92      0.361
## Year2008          0.5978      0.2846      2.10      0.038 *
## Year2009          0.3347      0.3040      1.10      0.273
## Year2010          0.2846      0.4005      0.71      0.479
## Year2011          0.4223      0.2976      1.42      0.159
## Year2012          0.4290      0.2921      1.47      0.145
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.321, Adjusted R-squared:  0.192
## Convergence in 37 IRWLS iterations
##
## Robustness weights:
## observation 11 is an outlier with |weight| = 0 ( < 0.00072);
## 14 weights are ~= 1. The remaining 123 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.212  0.889   0.945   0.900   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          7.25e-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 2.097 1          1.448
## LastAuthorFemale  1.960 1          1.400
## Year              3.651 16          1.041

```

Residuals from first and last author



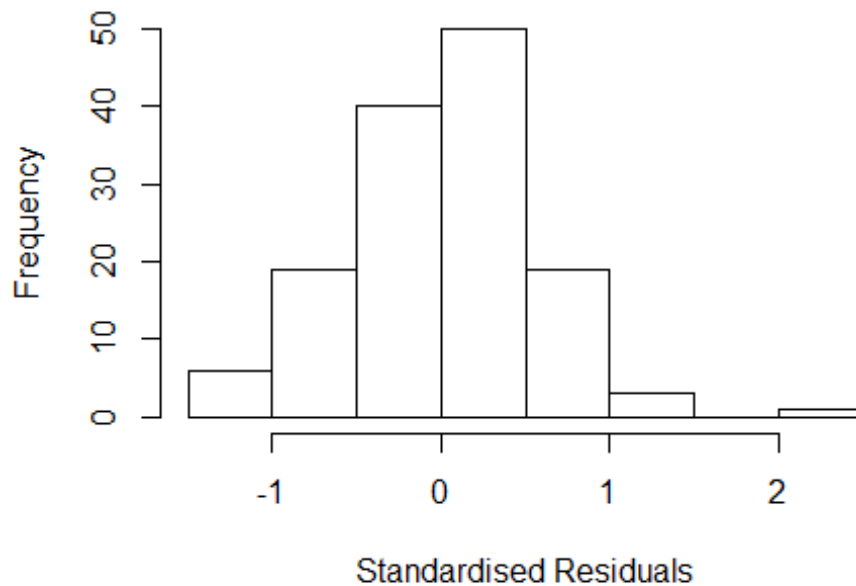
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3965 -0.2986 0.0292 0.3041 2.0080
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.78713 0.27781 2.83 0.0054 **
## FirstAuthorFemale1 -0.08085 0.10957 -0.74 0.4620
## LastAuthorFemale1 0.12950 0.10629 1.22 0.2255
## Year1997 -0.73315 0.31313 -2.34 0.0209 *
## Year1998 0.60939 0.89154 0.68 0.4956
## Year1999 -0.03133 0.41532 -0.08 0.9400
## Year2000 0.31572 0.36317 0.87 0.3864
## Year2001 0.27814 0.39337 0.71 0.4809
## Year2002 0.37993 0.28864 1.32 0.1906
## Year2003 0.00395 0.38013 0.01 0.9917
## Year2004 0.52233 0.29522 1.77 0.0794 .
## Year2005 0.16089 0.36854 0.44 0.6632
```

```

## Year2006          0.36352      0.33304      1.09      0.2772
## Year2007          0.21985      0.31152      0.71      0.4817
## Year2008          0.58103      0.28453      2.04      0.0434 *
## Year2009          0.27195      0.29872      0.91      0.3645
## Year2010          0.35033      0.33913      1.03      0.3037
## Year2011          0.39490      0.30487      1.30      0.1977
## Year2012          0.42879      0.28354      1.51      0.1331
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.523
## Multiple R-squared:  0.198, Adjusted R-squared:  0.0769
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 121 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##   0.107  0.836  0.963  0.896  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      7.25e-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 2.204 1      1.485
## Year              2.204 16      1.025

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3834 -0.3192 0.0314 0.2802 2.0161
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.83052 0.28090 2.96 0.0037 **
## FirstAuthorFemale1 -0.05280 0.10972 -0.48 0.6313
## Year1997 -0.78463 0.30979 -2.53 0.0126 *
## Year1998 0.55290 0.93753 0.59 0.5565
## Year1999 -0.04190 0.42690 -0.10 0.9220
## Year2000 0.31979 0.38320 0.83 0.4056
## Year2001 0.22411 0.39235 0.57 0.5689
## Year2002 0.35785 0.29449 1.22 0.2267
## Year2003 0.00216 0.39491 0.01 0.9957
## Year2004 0.46842 0.29359 1.60 0.1132
## Year2005 0.13232 0.38040 0.35 0.7286
## Year2006 0.35978 0.33863 1.06 0.2902
```

```

## Year2007          0.21996    0.32332    0.68    0.4976
## Year2008          0.58587    0.29354    2.00    0.0482 *
## Year2009          0.24207    0.30649    0.79    0.4312
## Year2010          0.34562    0.35572    0.97    0.3332
## Year2011          0.38020    0.30951    1.23    0.2217
## Year2012          0.41483    0.29179    1.42    0.1577
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.515
## Multiple R-squared:  0.192, Adjusted R-squared:  0.0771
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 122 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.091  0.849   0.959   0.893   0.991   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      7.25e-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 1.973  1      1.405
## Year              1.973 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3333 -0.2849  0.0445  0.3024  2.0118

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7707    0.2666   2.89  0.0046 **
## LastAuthorFemale1 0.1100    0.1074   1.02  0.3076
## Year1997       -0.7205    0.3022  -2.38  0.0187 *
## Year1998        0.5547    0.9041   0.61  0.5407
## Year1999       -0.0572    0.4123  -0.14  0.8898
## Year2000        0.3262    0.3678   0.89  0.3769
## Year2001        0.2674    0.3815   0.70  0.4848
## Year2002        0.3495    0.2761   1.27  0.2080
## Year2003        0.0116    0.3778   0.03  0.9756
## Year2004        0.5077    0.2827   1.80  0.0750 .
## Year2005        0.1531    0.3593   0.43  0.6709
## Year2006        0.3386    0.3212   1.05  0.2940
## Year2007        0.2201    0.3031   0.73  0.4691
## Year2008        0.5626    0.2741   2.05  0.0423 *
## Year2009        0.2564    0.2851   0.90  0.3704
## Year2010        0.3438    0.3261   1.05  0.2940
## Year2011        0.3901    0.2986   1.31  0.1939
## Year2012        0.4170    0.2710   1.54  0.1265
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.197, Adjusted R-squared:  0.0836
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 122 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0997 0.8500 0.9590 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      7.25e-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: 138"
## [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
## 27 35 2 3 4 30 1 19 28 19 26 27 29 27 34
## 2011 2012
## 35 31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 17 28 1 3 3 14 1 18 19 13 19 21 23 17 23
## 2011 2012
## 22 23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 15 22 1 3 3 13 1 17 18 10 18 20 22 17 23
## 2011 2012
## 21 21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.67118750558929"
## [1] "Male first author team size 2018 geometric mean: 1.66827985773183"

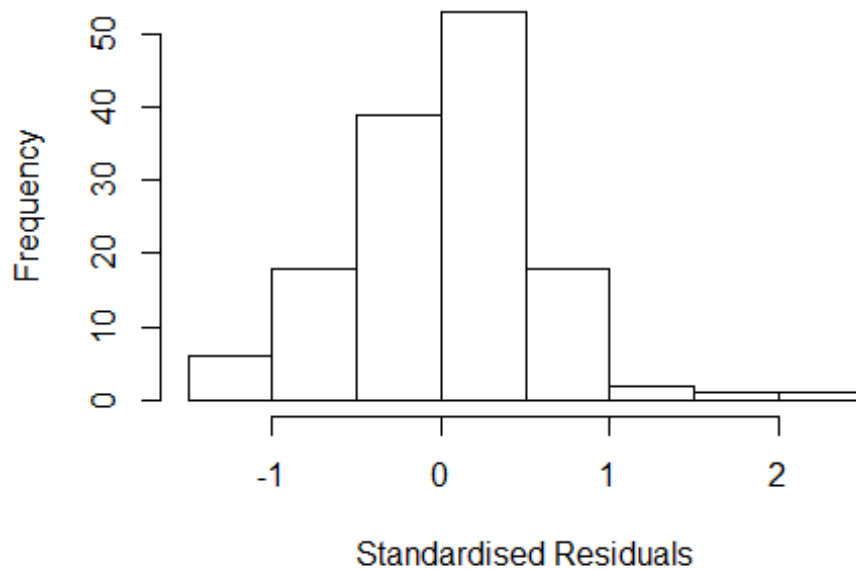
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 47, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.55970759249787"
## [1] "Male last author team size 2018 geometric mean: 1.81712059283214"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

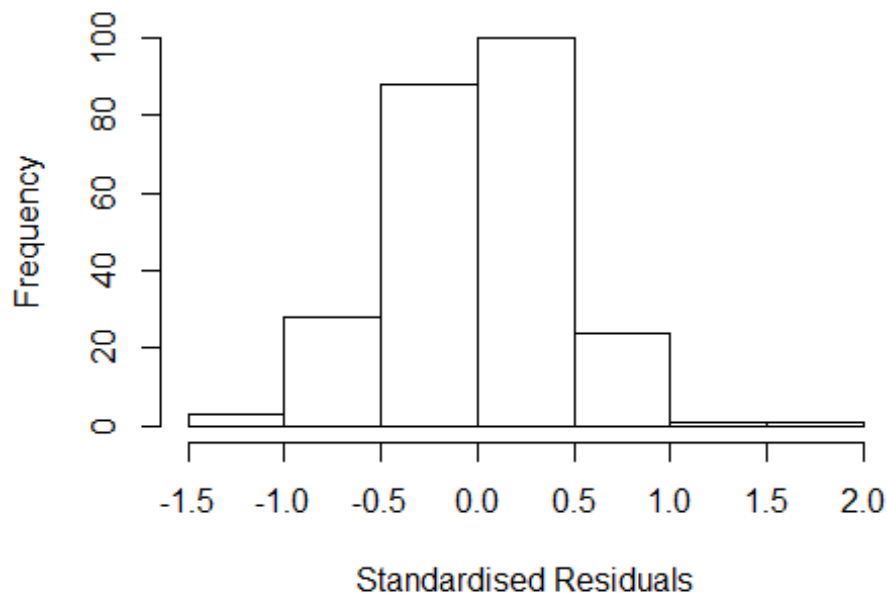
```


Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 42, 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 2.755 1      1.660
## LastAuthorFemale  3.357 1      1.832
## UniqueAuthors    36.157 4      1.566
## Year              53.409 16     1.132
```

Residuals from first and last author and team size



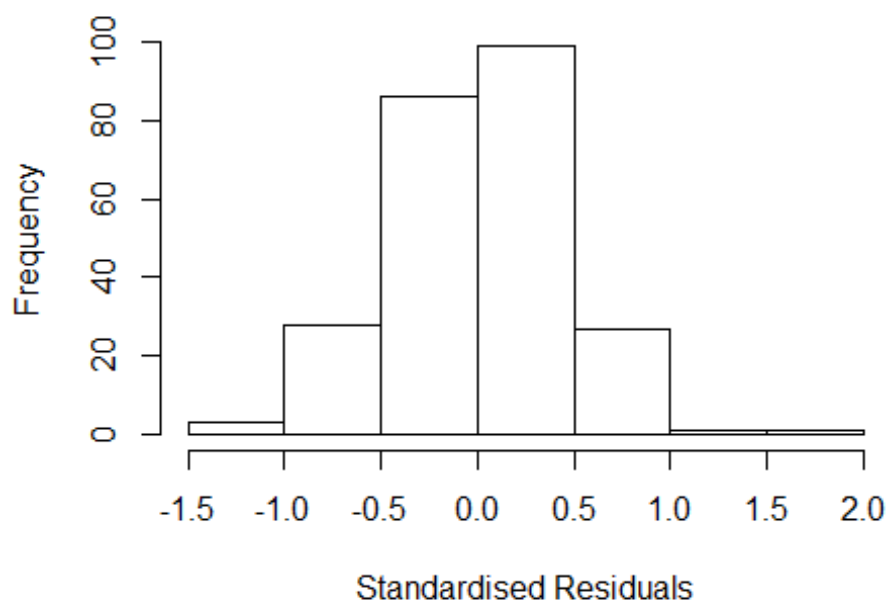
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1249 -0.2644 0.0195 0.2867 1.8034
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.99534 0.10902 9.13 < 2e-16 ***
## FirstAuthorFemale1 0.16985 0.07570 2.24 0.02584 *
## LastAuthorFemale1 -0.08766 0.07488 -1.17 0.24296
## UniqueAuthors2 0.04032 0.07485 0.54 0.59066
## UniqueAuthors3 0.01530 0.08774 0.17 0.86170
## UniqueAuthors4 0.01248 0.11067 0.11 0.91029
## UniqueAuthors5 0.00175 0.11847 0.01 0.98820
## Year1997 -0.00881 0.11894 -0.07 0.94101
## Year1998 -0.08609 0.15294 -0.56 0.57406
## Year1999 0.55336 0.35892 1.54 0.12457
```

```

## Year2000      -0.36862    0.21066   -1.75  0.08153 .
## Year2001      -0.21142    0.17232   -1.23  0.22116
## Year2002      -0.57509    0.15294   -3.76  0.00022 ***
## Year2003      -0.11533    0.15677   -0.74  0.46269
## Year2004      -0.04883    0.17127   -0.29  0.77583
## Year2005      -0.12132    0.15874   -0.76  0.44553
## Year2006       0.10109    0.14127    0.72  0.47498
## Year2007      -0.20174    0.19371   -1.04  0.29878
## Year2008       0.02056    0.13057    0.16  0.87503
## Year2009      -0.04200    0.24041   -0.17  0.86148
## Year2010       0.05261    0.13079    0.40  0.68789
## Year2011      -0.06266    0.12961   -0.48  0.62924
## Year2012      -0.06547    0.13991   -0.47  0.64029
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.104, Adjusted R-squared:  0.0152
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 218 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0301 0.8710 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      4.08e-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 2.764 1 1.663
## LastAuthorFemale 1.886 1 1.373
## Year 2.332 16 1.027

```

Residuals from first and last author



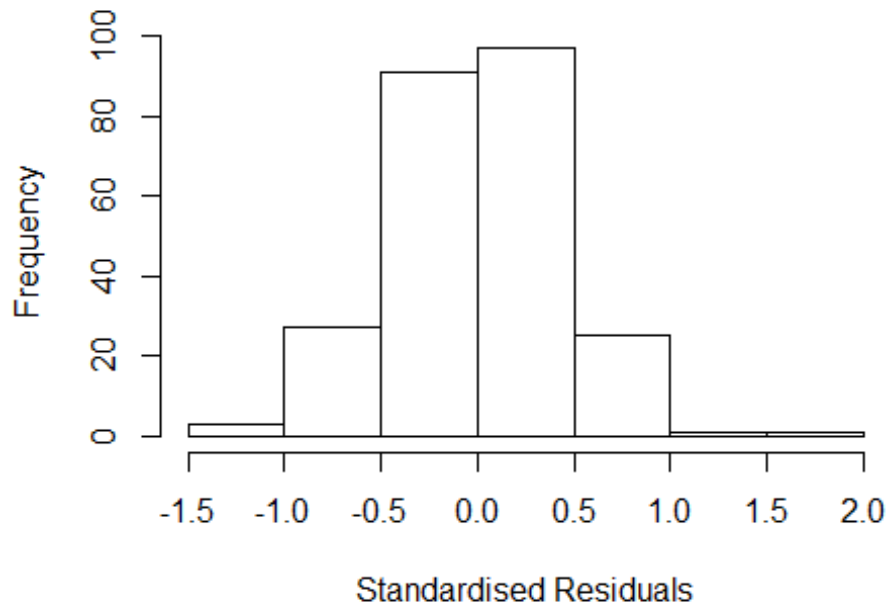
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1377 -0.2637 0.0172 0.2773 1.7852
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.01034 0.10515 9.61 < 2e-16 ***
## FirstAuthorFemale1 0.16996 0.07336 2.32 0.021 *
## LastAuthorFemale1 -0.09024 0.07276 -1.24 0.216
## Year1997 -0.00398 0.12038 -0.03 0.974
## Year1998 -0.09934 0.10515 -0.94 0.346
## Year1999 0.54134 0.33445 1.62 0.107
## Year2000 -0.37981 0.21065 -1.80 0.073 .
## Year2001 -0.21636 0.16897 -1.28 0.202
## Year2002 -0.58834 0.10515 -5.60 6.3e-08 ***
## Year2003 -0.11913 0.15675 -0.76 0.448
## Year2004 -0.04492 0.17009 -0.26 0.792
## Year2005 -0.11512 0.16014 -0.72 0.473
```

```

## Year2006          0.10449    0.13975    0.75    0.455
## Year2007          -0.19849    0.18038   -1.10    0.272
## Year2008           0.02132    0.13142    0.16    0.871
## Year2009          -0.04263    0.24350   -0.18    0.861
## Year2010           0.05491    0.13154    0.42    0.677
## Year2011          -0.06441    0.12555   -0.51    0.608
## Year2012          -0.05964    0.13748   -0.43    0.665
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.425
## Multiple R-squared:  0.103, Adjusted R-squared:  0.0313
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 219 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0389 0.8710 0.9460 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      4.08e-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 7.058 1      2.657
## Year              7.058 16      1.063

```

Residuals from first author

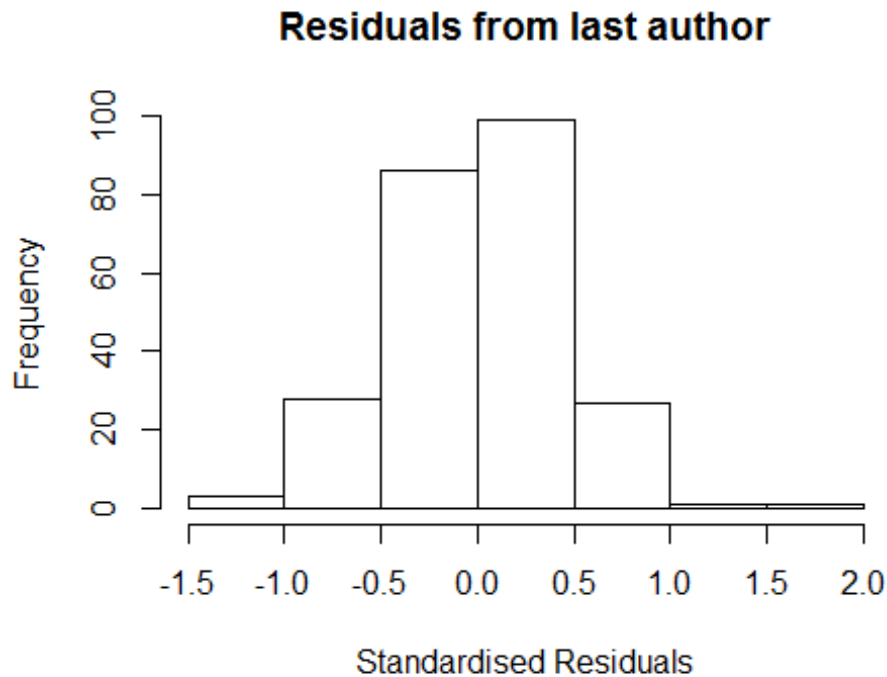


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.15099 -0.30147 0.00881 0.29439 1.79203
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.99534 0.10572 9.41 < 2e-16 ***
## FirstAuthorFemale1 0.13060 0.06348 2.06 0.041 *
## Year1997 0.00485 0.12301 0.04 0.969
## Year1998 -0.08434 0.10572 -0.80 0.426
## Year1999 0.51825 0.33785 1.53 0.126
## Year2000 -0.36471 0.21031 -1.73 0.084 .
## Year2001 -0.21601 0.16734 -1.29 0.198
## Year2002 -0.57334 0.10572 -5.42 1.5e-07 ***
## Year2003 -0.10774 0.15641 -0.69 0.492
## Year2004 -0.05024 0.17109 -0.29 0.769
## Year2005 -0.10508 0.15821 -0.66 0.507
## Year2006 0.09801 0.14148 0.69 0.489
```

```

## Year2007      -0.19037    0.18291   -1.04    0.299
## Year2008      0.02505    0.13270    0.19    0.850
## Year2009     -0.03505    0.23147   -0.15    0.880
## Year2010      0.06219    0.13245    0.47    0.639
## Year2011     -0.05847    0.12681   -0.46    0.645
## Year2012     -0.05402    0.13868   -0.39    0.697
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.093, Adjusted R-squared:  0.0251
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 223 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0561 0.8810 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      4.08e-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 1.55 1          1.245
## Year            1.55 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.15446 -0.28846 0.00746 0.28825 1.75909
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.02e+00 1.07e-01 9.51 <2e-16 ***
## LastAuthorFemale1 -1.39e-02 6.37e-02 -0.22 0.827
## Year1997 2.99e-02 1.26e-01 0.24 0.813
## Year1998 -1.08e-01 1.07e-01 -1.01 0.313
## Year1999 4.99e-01 3.39e-01 1.47 0.142
## Year2000 -3.89e-01 2.11e-01 -1.84 0.067 .
## Year2001 -2.12e-01 1.66e-01 -1.28 0.201
## Year2002 -5.97e-01 1.07e-01 -5.58 7e-08 ***
## Year2003 -9.61e-02 1.61e-01 -0.60 0.550
## Year2004 -3.28e-02 1.74e-01 -0.19 0.850
## Year2005 -6.75e-02 1.54e-01 -0.44 0.661
## Year2006 1.35e-01 1.43e-01 0.95 0.344
```

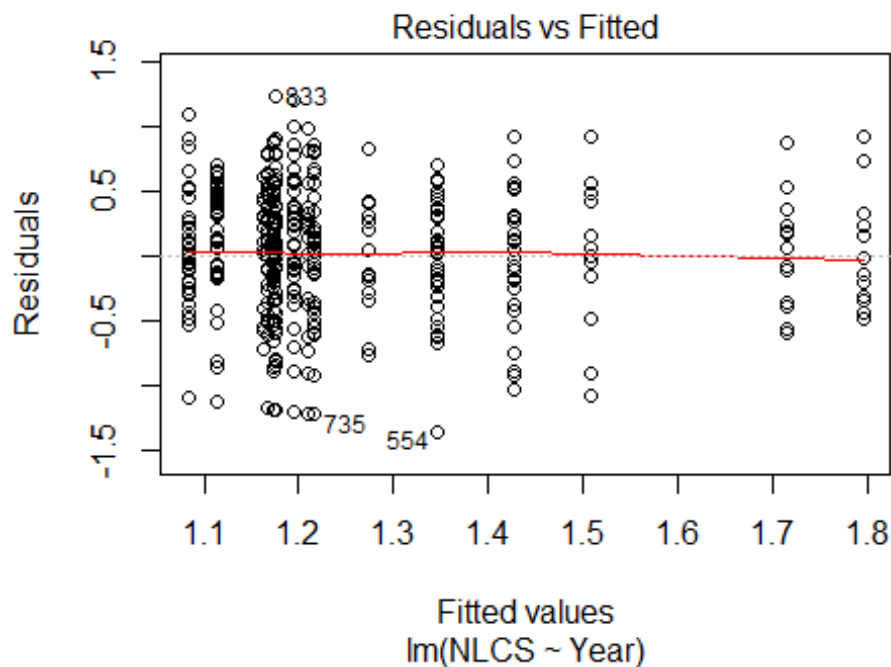


```

## Year2007          -1.82e-01   1.87e-01   -0.97   0.333
## Year2008           4.92e-02   1.32e-01    0.37   0.709
## Year2009          -3.83e-02   2.38e-01   -0.16   0.872
## Year2010           1.23e-01   1.33e-01    0.93   0.354
## Year2011           9.63e-05   1.27e-01    0.00   0.999
## Year2012           2.31e-02   1.38e-01    0.17   0.867
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.433
## Multiple R-squared:  0.0765, Adjusted R-squared:  0.00738
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 218 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0612 0.8620 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      4.08e-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: 245"
## [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
##   26   37   37   30   47   42   40   41   36   51   54   53   51   47   46
## 2011 2012
##   44   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   16   13   19   20   12   17   15   15   22   39   41   44   39   36
## 2011 2012

```

```
## 36 44
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 15 10 17 18 11 14 14 14 19 32 38 37 32 34
## 2011 2012
## 35 40
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 8.8, df = 16, p-value = 0.9
```



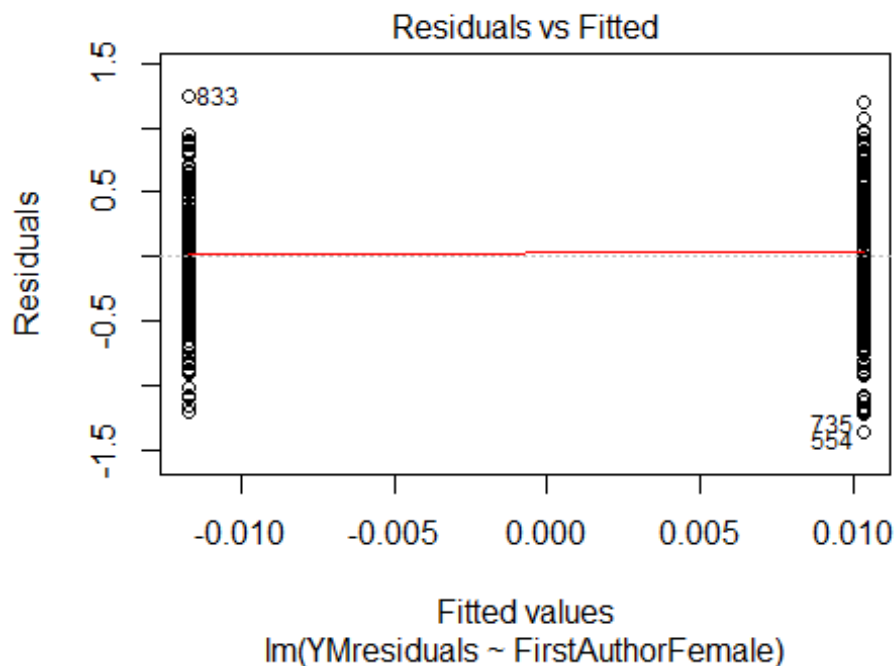
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 5.12441921150926"
## [1] "Male first author team size 2018 geometric mean: 4.28216661910156"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

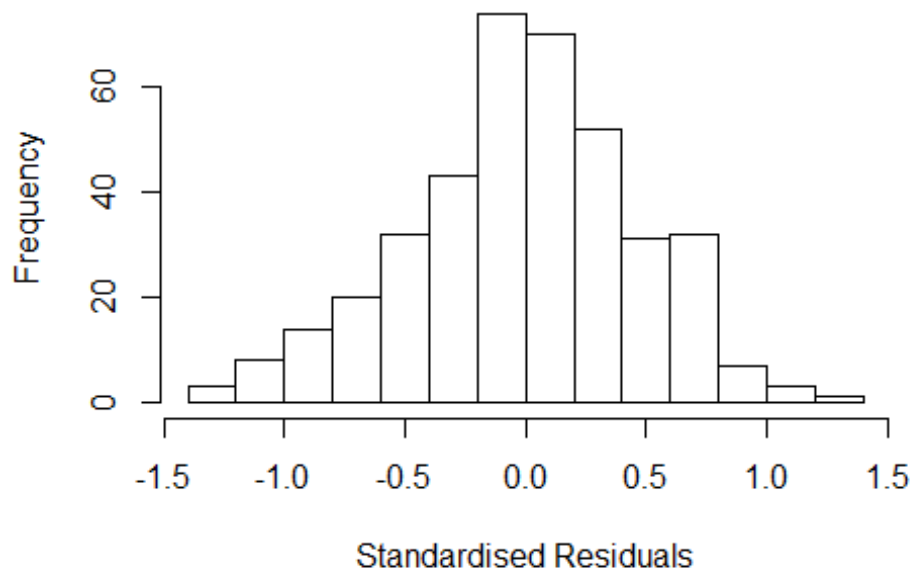
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 250, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.28482662483017"
## [1] "Male last author team size 2018 geometric mean: 5.08547154606966"

## 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 = 0.1
## 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 1.344 1      1.159
## LastAuthorFemale  1.281 1      1.132
## UniqueAuthors    3.010 4      1.148
## Year              3.626 16     1.041
```

Residuals from first and last author and team size



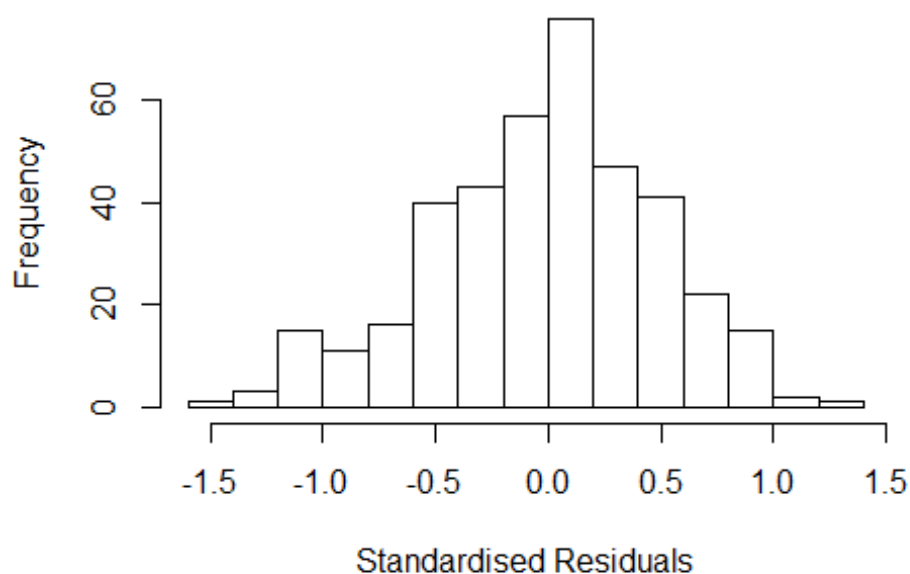
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.38728 -0.29541 0.00433 0.31233 1.27246
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1804 0.1923 6.14 2.2e-09 ***
## FirstAuthorFemale1 -0.0832 0.0537 -1.55 0.12210
## LastAuthorFemale1 0.0675 0.0520 1.30 0.19500
## UniqueAuthors2 0.2469 0.0964 2.56 0.01086 *
## UniqueAuthors3 0.3280 0.0987 3.32 0.00098 ***
## UniqueAuthors4 0.4923 0.0930 5.30 2.1e-07 ***
## UniqueAuthors5 0.4917 0.0824 5.96 5.8e-09 ***
## Year1997 0.2523 0.2322 1.09 0.27791
## Year1998 0.2911 0.2550 1.14 0.25445
## Year1999 -0.3169 0.2350 -1.35 0.17828
```

```

## Year2000          -0.0678      0.2058   -0.33   0.74199
## Year2001           0.0890      0.2782    0.32   0.74909
## Year2002          -0.4685      0.2174   -2.15   0.03184 *
## Year2003          -0.2888      0.2179   -1.33   0.18574
## Year2004          -0.3720      0.2131   -1.75   0.08170 .
## Year2005          -0.3592      0.2192   -1.64   0.10225
## Year2006          -0.3736      0.2145   -1.74   0.08232 .
## Year2007          -0.1139      0.2033   -0.56   0.57564
## Year2008          -0.4105      0.2152   -1.91   0.05727 .
## Year2009          -0.3033      0.2067   -1.47   0.14303
## Year2010          -0.2808      0.2129   -1.32   0.18793
## Year2011          -0.3206      0.2086   -1.54   0.12524
## Year2012          -0.3583      0.2087   -1.72   0.08684 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.467
## Multiple R-squared:  0.207, Adjusted R-squared:  0.16
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 349 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.357  0.855   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           2.56e-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 1.252 1           1.119
## LastAuthorFemale  1.251 1           1.119
## Year              1.564 16           1.014

```

Residuals from first and last author



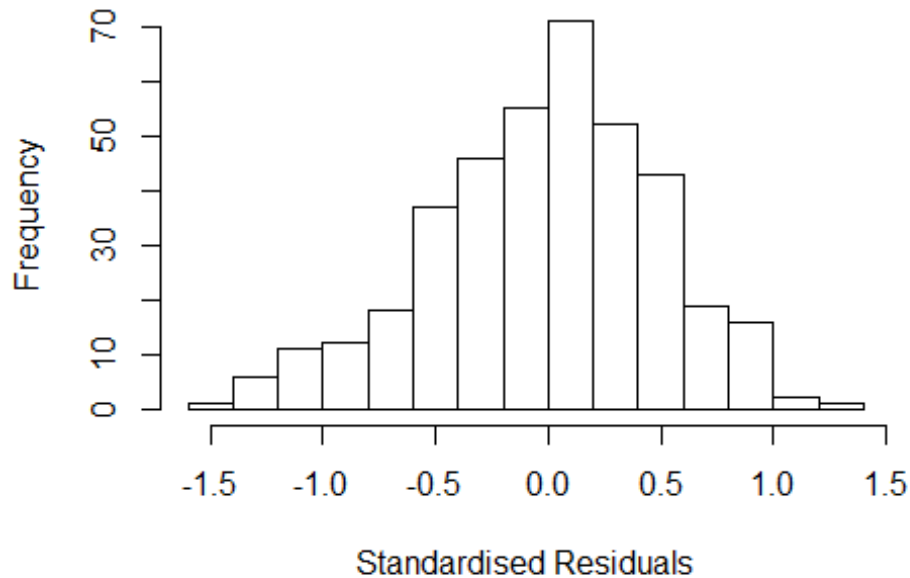
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4064 -0.3401  0.0184  0.3220  1.2350
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.41666    0.19608    7.22 2.9e-12 ***
## FirstAuthorFemale1 -0.04690    0.05664   -0.83  0.41
## LastAuthorFemale1  0.05694    0.05574    1.02  0.31
## Year1997          0.27709    0.23313    1.19  0.24
## Year1998          0.34917    0.24541    1.42  0.16
## Year1999         -0.24357    0.24278   -1.00  0.32
## Year2000         -0.00188    0.22201   -0.01  0.99
## Year2001          0.06988    0.27563    0.25  0.80
## Year2002         -0.31997    0.22591   -1.42  0.16
## Year2003         -0.15290    0.23136   -0.66  0.51
## Year2004         -0.27686    0.23298   -1.19  0.24
## Year2005         -0.35225    0.23347   -1.51  0.13
```

```

## Year2006      -0.23815    0.22857   -1.04    0.30
## Year2007      -0.01025    0.20762   -0.05    0.96
## Year2008      -0.28426    0.21943   -1.30    0.20
## Year2009      -0.21239    0.21380   -0.99    0.32
## Year2010      -0.17281    0.21734   -0.80    0.43
## Year2011      -0.24064    0.21513   -1.12    0.26
## Year2012      -0.25471    0.21642   -1.18    0.24
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.49
## Multiple R-squared:  0.101, Adjusted R-squared:  0.0578
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 39 weights are ~= 1. The remaining 351 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.389  0.870  0.946  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      2.56e-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 1.268 1      1.126
## Year              1.268 16      1.007

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.4157 -0.3211 0.0148 0.3131 1.2632
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.43342 0.19521 7.34 1.3e-12 ***
## FirstAuthorFemale1 -0.04053 0.05721 -0.71 0.48
## Year1997 0.27902 0.23223 1.20 0.23
## Year1998 0.34338 0.24739 1.39 0.17
## Year1999 -0.23298 0.24362 -0.96 0.34
## Year2000 0.00663 0.22146 0.03 0.98
## Year2001 0.06716 0.27286 0.25 0.81
## Year2002 -0.30670 0.22629 -1.36 0.18
## Year2003 -0.16343 0.22789 -0.72 0.47
## Year2004 -0.27291 0.23262 -1.17 0.24
## Year2005 -0.35303 0.23392 -1.51 0.13
## Year2006 -0.23324 0.22887 -1.02 0.31
```

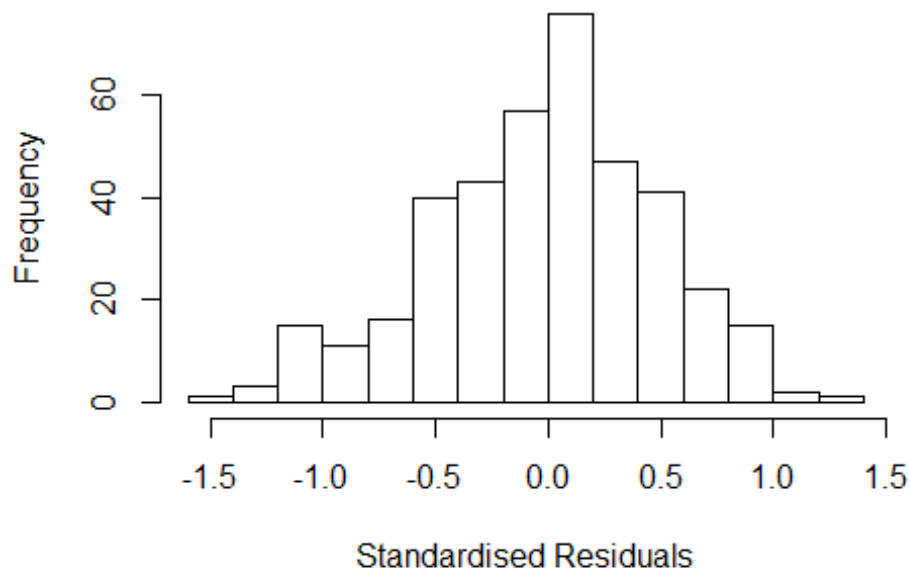


```

## Year2007          -0.01777    0.20773   -0.09    0.93
## Year2008          -0.28724    0.22038   -1.30    0.19
## Year2009          -0.20881    0.21405   -0.98    0.33
## Year2010          -0.17070    0.21824   -0.78    0.43
## Year2011          -0.23617    0.21493   -1.10    0.27
## Year2012          -0.24907    0.21612   -1.15    0.25
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.488
## Multiple R-squared:  0.0988, Adjusted R-squared:  0.0576
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 39 weights are ~= 1. The remaining 351 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.381  0.869  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      2.56e-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 1.239 1          1.113
## Year            1.239 16          1.007

```

Residuals from last author



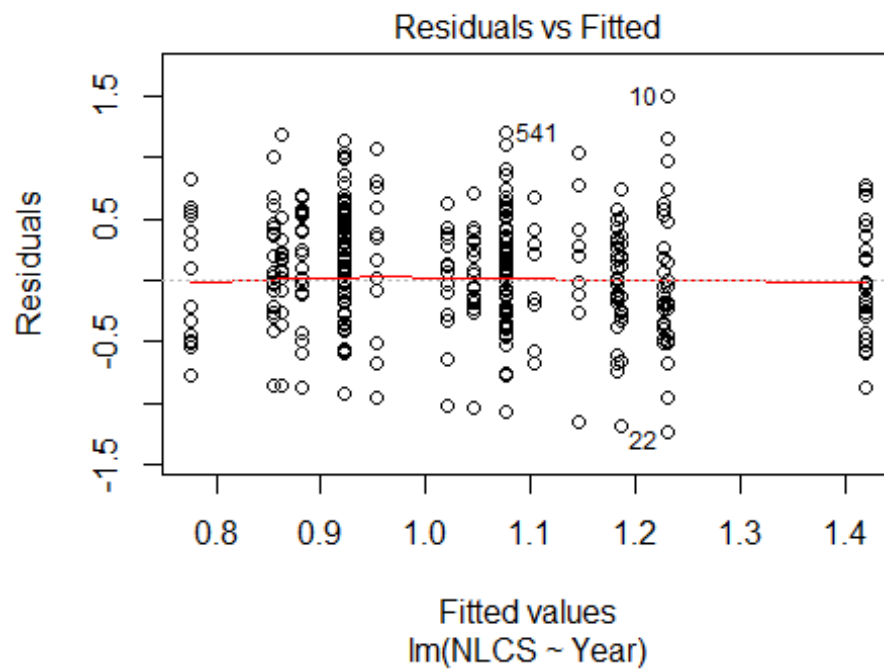
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3858 -0.3432  0.0224  0.3393  1.2127
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4104     0.1974   7.15 4.8e-12 ***
## LastAuthorFemale1  0.0515     0.0563   0.91  0.36
## Year1997          0.2714     0.2348   1.16  0.25
## Year1998          0.3410     0.2491   1.37  0.17
## Year1999         -0.2651     0.2426  -1.09  0.28
## Year2000         -0.0217     0.2224  -0.10  0.92
## Year2001          0.0708     0.2758   0.26  0.80
## Year2002         -0.3345     0.2248  -1.49  0.14
## Year2003         -0.1637     0.2322  -0.71  0.48
## Year2004         -0.2799     0.2358  -1.19  0.24
## Year2005         -0.3567     0.2342  -1.52  0.13
## Year2006         -0.2620     0.2246  -1.17  0.24
```

```

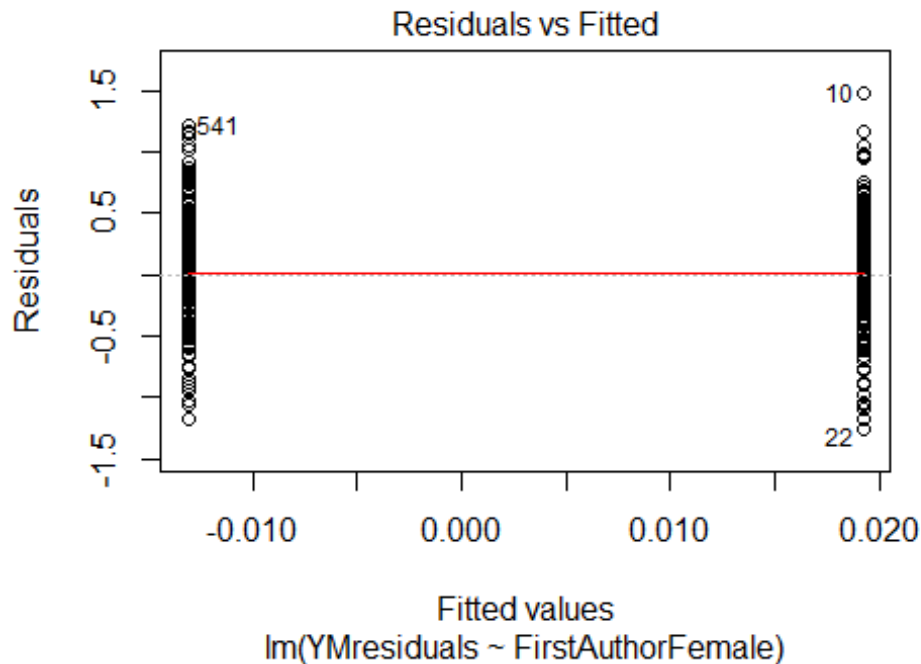
## Year2007          -0.0246      0.2077    -0.12      0.91
## Year2008          -0.3061      0.2178    -1.41      0.16
## Year2009          -0.2355      0.2117    -1.11      0.27
## Year2010          -0.1838      0.2175    -0.84      0.40
## Year2011          -0.2612      0.2140    -1.22      0.22
## Year2012          -0.2676      0.2178    -1.23      0.22
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.0996, Adjusted R-squared:  0.0584
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 357 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.402  0.870   0.951   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      2.56e-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: 390"
## [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
##   24   14   22   36   24   31   19   21   27   29   16   24   20   30   29
## 2011 2012
##   90  115
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   17    6   10   28   11    8   14   16   25   25   14   17   14   24   19
## 2011 2012

```

```
## 76 93
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 13 6 9 28 8 8 13 16 24 23 13 15 13 24 19
## 2011 2012
## 72 84
## [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
```

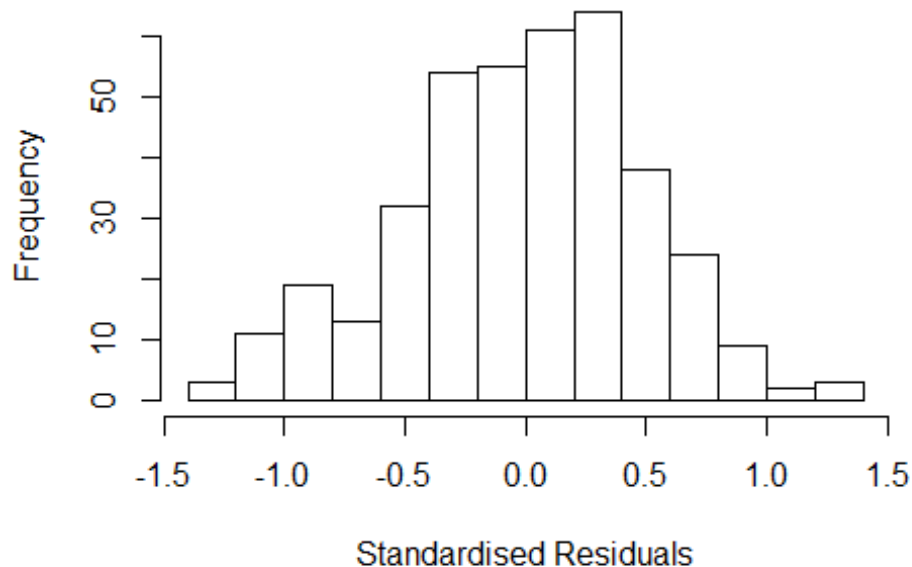


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.97, df = 1, p-value = 0.3
```



```
## [1] "Female first author team size 2018 geometric mean: 1.62305041177008"
## [1] "Male first author team size 2018 geometric mean: 1.69003029417259"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.5261901233615"
## [1] "Male last author team size 2018 geometric mean: 1.8563162782099"
##
## 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 1.605 1      1.267
## LastAuthorFemale  1.539 1      1.241
## UniqueAuthors     2.339 4      1.112
## Year              3.432 16     1.039
```

Residuals from first and last author and team size



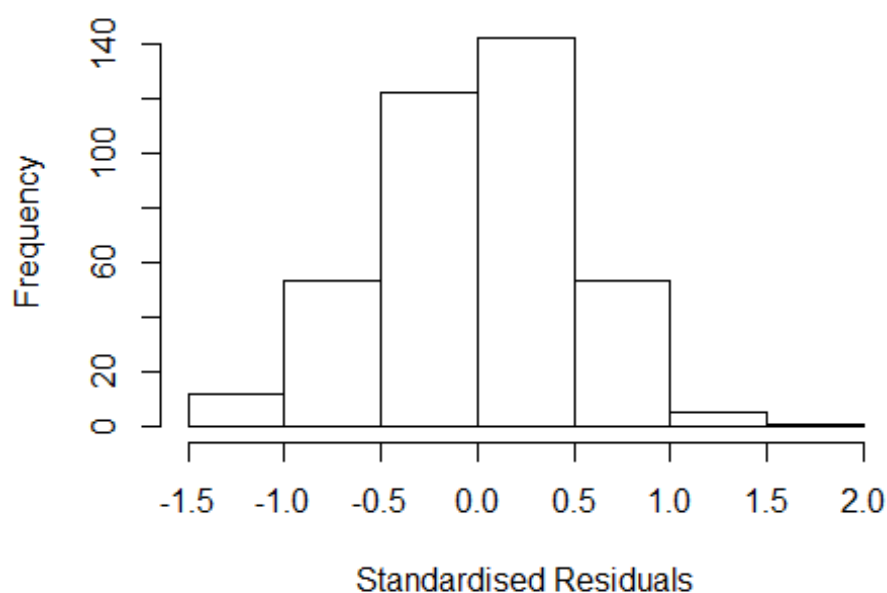
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2817 -0.3229 0.0285 0.3177 1.3669
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2359 0.2186 5.65 3.2e-08 ***
## FirstAuthorFemale1 0.0389 0.0630 0.62 0.53745
## LastAuthorFemale1 -0.0476 0.0599 -0.79 0.42752
## UniqueAuthors2 0.0184 0.0785 0.23 0.81480
## UniqueAuthors3 0.1728 0.1017 1.70 0.09006 .
## UniqueAuthors4 0.2675 0.0853 3.14 0.00186 **
## UniqueAuthors5 0.2718 0.0818 3.32 0.00098 ***
## Year1997 -0.0831 0.2569 -0.32 0.74660
## Year1998 0.0458 0.3445 0.13 0.89435
## Year1999 0.0722 0.2338 0.31 0.75769
```

```

## Year2000          -0.0798      0.2649   -0.30   0.76344
## Year2001          -0.2254      0.2594   -0.87   0.38539
## Year2002          -0.3693      0.2597   -1.42   0.15580
## Year2003          -0.2251      0.2464   -0.91   0.36158
## Year2004          -0.4785      0.2396   -2.00   0.04655 *
## Year2005          -0.4209      0.2428   -1.73   0.08383 .
## Year2006          -0.6477      0.2711   -2.39   0.01739 *
## Year2007          -0.5319      0.2597   -2.05   0.04128 *
## Year2008          -0.5198      0.3099   -1.68   0.09427 .
## Year2009          -0.2231      0.2339   -0.95   0.34099
## Year2010          -0.3274      0.2389   -1.37   0.17140
## Year2011          -0.2047      0.2311   -0.89   0.37618
## Year2012          -0.3683      0.2308   -1.60   0.11140
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.492
## Multiple R-squared:  0.149, Adjusted R-squared:  0.0979
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 362 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.421  0.878  0.956  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          2.58e-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 1.685 1          1.298
## LastAuthorFemale  1.546 1          1.243
## Year              1.603 16          1.015

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.3104 -0.3195 0.0139 0.3590 1.5126
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3240 0.2306 5.74 2e-08 ***
## FirstAuthorFemale1 0.0467 0.0654 0.71 0.475
## LastAuthorFemale1 -0.1086 0.0625 -1.74 0.083 .
## Year1997 -0.0785 0.2808 -0.28 0.780
## Year1998 -0.0136 0.3586 -0.04 0.970
## Year1999 0.1073 0.2477 0.43 0.665
## Year2000 0.0289 0.2754 0.10 0.916
## Year2001 -0.2030 0.2878 -0.71 0.481
## Year2002 -0.3161 0.2663 -1.19 0.236
## Year2003 -0.0865 0.2623 -0.33 0.742
## Year2004 -0.4323 0.2590 -1.67 0.096 .
## Year2005 -0.3897 0.2622 -1.49 0.138
```

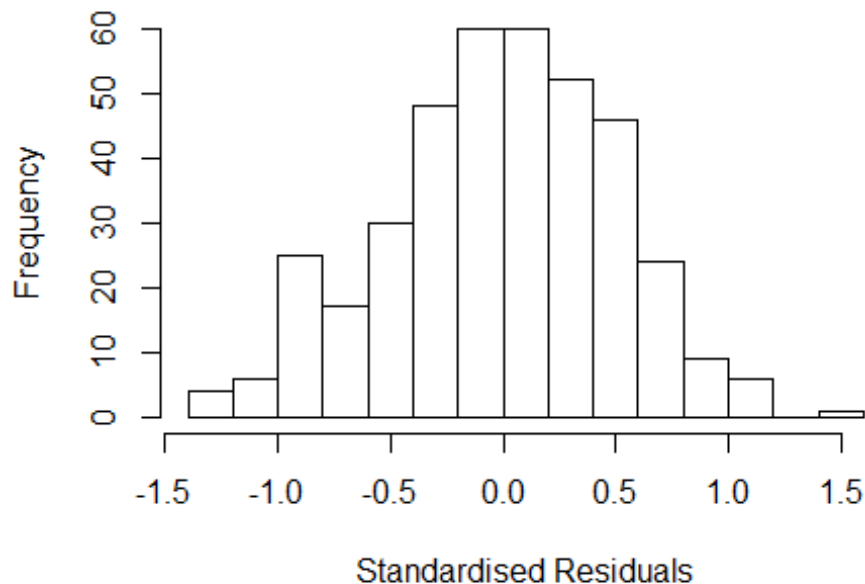


```

## Year2006          -0.5841      0.2777    -2.10      0.036 *
## Year2007          -0.4112      0.2724    -1.51      0.132
## Year2008          -0.4003      0.3184    -1.26      0.209
## Year2009          -0.1287      0.2498    -0.52      0.607
## Year2010          -0.2124      0.2562    -0.83      0.408
## Year2011          -0.1958      0.2445    -0.80      0.424
## Year2012          -0.3591      0.2433    -1.48      0.141
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.113, Adjusted R-squared:  0.07
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 364 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.328  0.877  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      2.58e-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 1.326 1      1.152
## Year              1.326 16      1.009

```

Residuals from first author



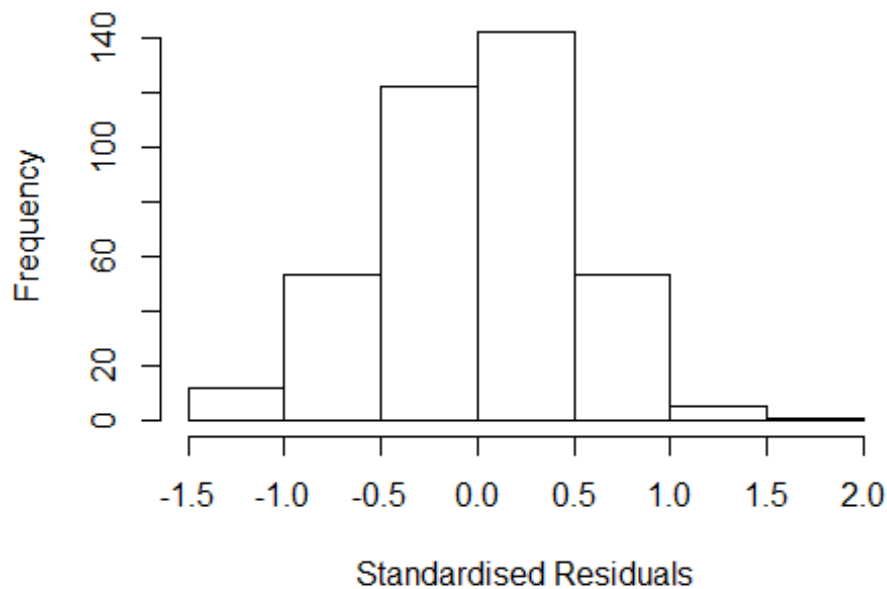
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2832 -0.3197 0.0132 0.3606 1.4600
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.268013 0.214852 5.90 8.1e-09 ***
## FirstAuthorFemale1 0.000227 0.058545 0.00 0.997
## Year1997 -0.044765 0.264061 -0.17 0.865
## Year1998 0.015186 0.346061 0.04 0.965
## Year1999 0.149196 0.234295 0.64 0.525
## Year2000 0.091660 0.262926 0.35 0.728
## Year2001 -0.158862 0.279964 -0.57 0.571
## Year2002 -0.256976 0.249739 -1.03 0.304
## Year2003 -0.050372 0.249600 -0.20 0.840
## Year2004 -0.421143 0.246828 -1.71 0.089 .
## Year2005 -0.362348 0.250000 -1.45 0.148
## Year2006 -0.548325 0.274679 -2.00 0.047 *
```

```

## Year2007          -0.395769    0.258933   -1.53    0.127
## Year2008          -0.377793    0.318430   -1.19    0.236
## Year2009          -0.078791    0.234337   -0.34    0.737
## Year2010          -0.187392    0.240231   -0.78    0.436
## Year2011          -0.177034    0.229996   -0.77    0.442
## Year2012          -0.337359    0.228926   -1.47    0.141
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.5
## Multiple R-squared:  0.106, Adjusted R-squared:  0.0652
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 364 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.374  0.871  0.950  0.908  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      2.58e-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 1.26 1          1.122
## Year              1.26 16          1.007

```

Residuals from last author



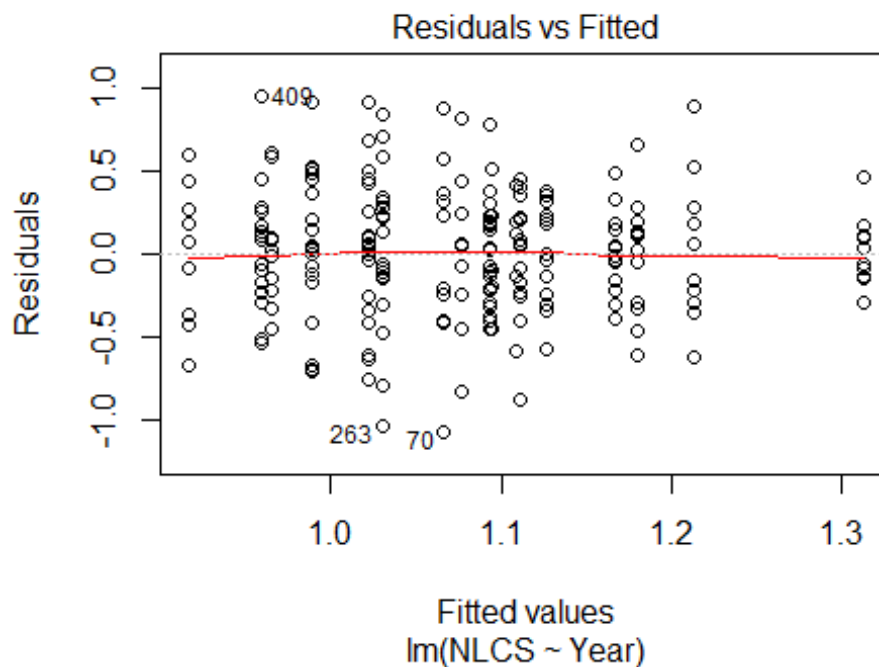
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3139 -0.3324 0.0207 0.3611 1.5036
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.314375 0.227962 5.77 1.7e-08 ***
## LastAuthorFemale1 -0.089969 0.056331 -1.60 0.111
## Year1997 -0.058889 0.278977 -0.21 0.833
## Year1998 -0.000453 0.355513 0.00 0.999
## Year1999 0.131980 0.242723 0.54 0.587
## Year2000 0.045243 0.270772 0.17 0.867
## Year2001 -0.183258 0.281979 -0.65 0.516
## Year2002 -0.289962 0.260216 -1.11 0.266
## Year2003 -0.055554 0.256275 -0.22 0.829
## Year2004 -0.398388 0.248886 -1.60 0.110
## Year2005 -0.353380 0.252763 -1.40 0.163
## Year2006 -0.548387 0.272893 -2.01 0.045 *
```

```

## Year2007      -0.394198    0.269391   -1.46    0.144
## Year2008      -0.375467    0.320179   -1.17    0.242
## Year2009      -0.093072    0.239840   -0.39    0.698
## Year2010      -0.186922    0.249492   -0.75    0.454
## Year2011      -0.166624    0.237900   -0.70    0.484
## Year2012      -0.331217    0.237853   -1.39    0.165
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.497
## Multiple R-squared:  0.112, Adjusted R-squared:  0.0711
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 360 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.339  0.876  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      2.58e-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: 388"
## [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
##    9   19   17   20   18   18   16   16   17   20   25   23   21   32   37
## 2011 2012
##   28   35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    9    9   10   11    5   11   11   13   13   19   15   14   22   18
## 2011 2012

```

```
## 23 23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 7 9 10 9 3 9 11 10 12 18 14 14 20 18
## 2011 2012
## 20 21
## [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
```



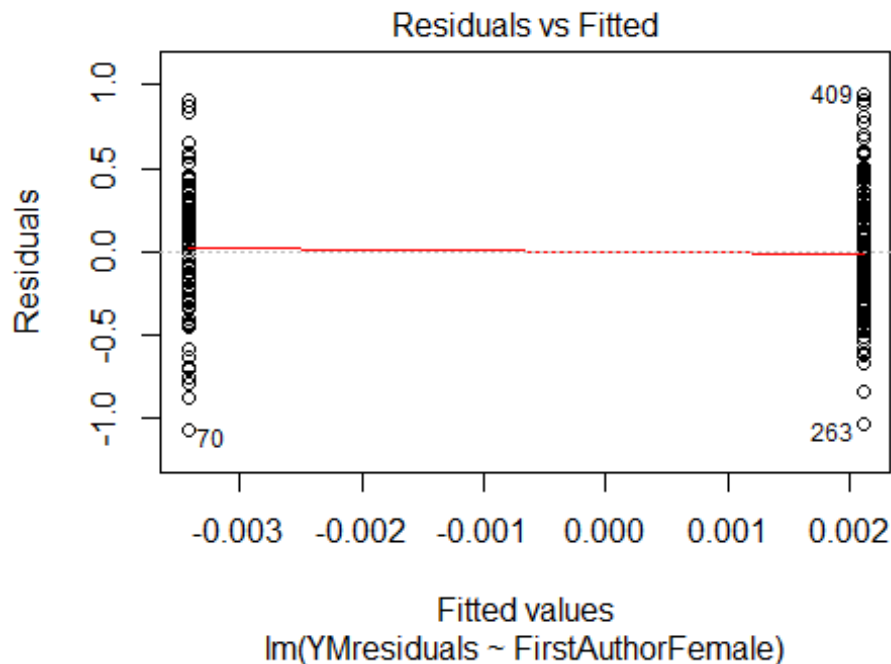
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.2, df = 1, p-value = 0.3

## [1] "Female first author team size 2018 geometric mean: 4.37564343898913"
## [1] "Male first author team size 2018 geometric mean: 3.89453837997481"

## 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 = 62, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.09062348923505"
## [1] "Male last author team size 2018 geometric mean: 4.08327071932403"

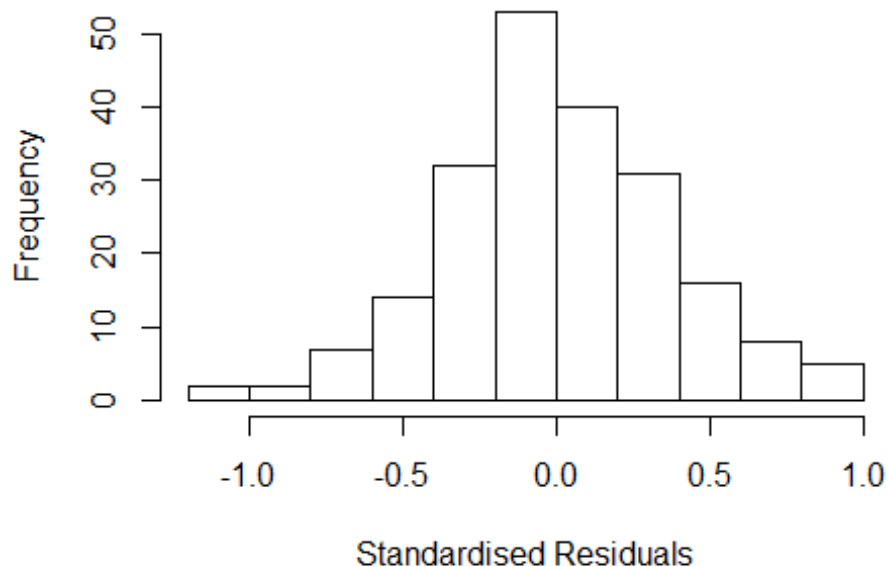
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 35, 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	1.388	1	1.178
LastAuthorFemale	1.902	1	1.379
UniqueAuthors	7.871	4	1.294
Year	9.587	16	1.073

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.10405 -0.20872 -0.00932 0.21869 0.96128
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.91466 0.26706 3.42 0.00076 ***
## FirstAuthorFemale1 -0.00157 0.05904 -0.03 0.97876
## LastAuthorFemale1 0.00308 0.06235 0.05 0.96068
## UniqueAuthors2 0.00331 0.19286 0.02 0.98633
## UniqueAuthors3 0.04917 0.18839 0.26 0.79436
## UniqueAuthors4 0.15075 0.18884 0.80 0.42571
## UniqueAuthors5 0.19705 0.18079 1.09 0.27715
## Year1997 -0.17108 0.28749 -0.60 0.55250
## Year1998 0.00876 0.25384 0.03 0.97252
## Year1999 0.14179 0.28139 0.50 0.61492
```

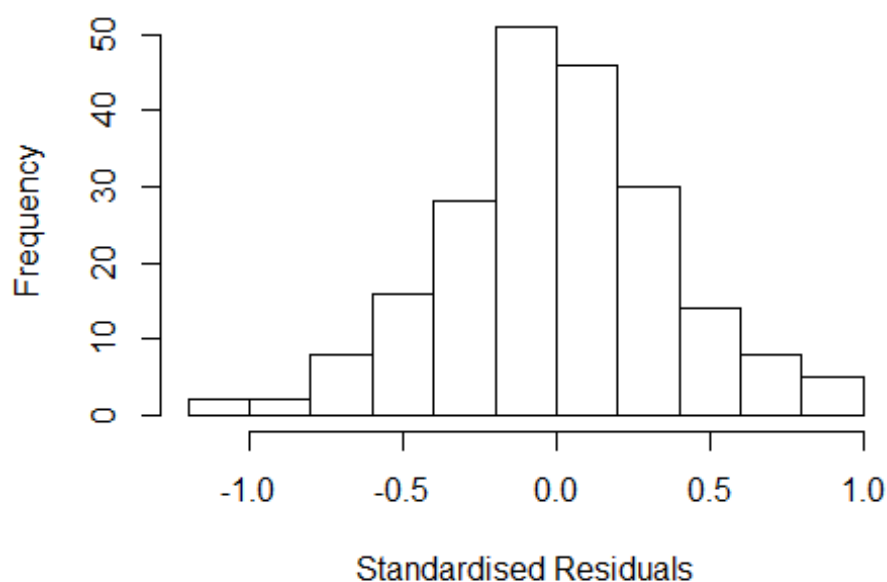


```

## Year2000      0.16628    0.28275    0.59  0.55718
## Year2001      0.18674    0.36569    0.51  0.61020
## Year2002     -0.10628    0.21070   -0.50  0.61455
## Year2003      0.26476    0.20719    1.28  0.20289
## Year2004      0.23680    0.21359    1.11  0.26900
## Year2005      0.11718    0.21948    0.53  0.59404
## Year2006     -0.01634    0.22286   -0.07  0.94162
## Year2007      0.13148    0.21247    0.62  0.53679
## Year2008      0.10622    0.20568    0.52  0.60617
## Year2009      0.02679    0.21371    0.13  0.90037
## Year2010     -0.11232    0.21479   -0.52  0.60165
## Year2011      0.03326    0.21006    0.16  0.87434
## Year2012      0.02614    0.21752    0.12  0.90448
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.345
## Multiple R-squared:  0.112, Adjusted R-squared:  0.00732
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 185 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.284  0.856  0.956  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      4.76e-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 1.361 1      1.167
## LastAuthorFemale  1.489 1      1.220
## Year              1.859 16      1.020

```

Residuals from first and last author



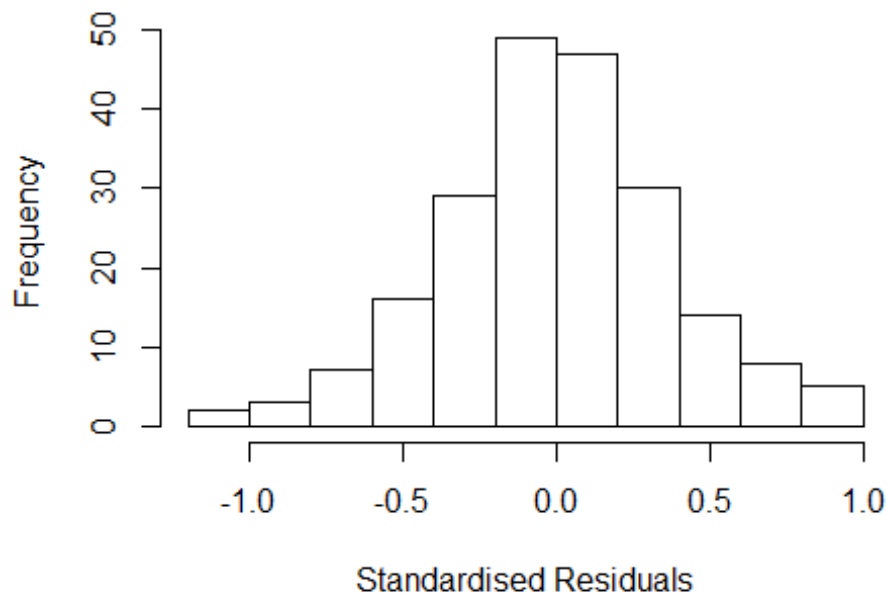
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.092 -0.231 -0.012 0.233 0.962
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09068 0.16302 6.69 2.4e-10 ***
## FirstAuthorFemale1 -0.02523 0.06181 -0.41 0.68
## LastAuthorFemale1 0.00863 0.05688 0.15 0.88
## Year1997 -0.25014 0.24859 -1.01 0.32
## Year1998 -0.00241 0.23337 -0.01 0.99
## Year1999 0.02623 0.27427 0.10 0.92
## Year2000 0.05065 0.23192 0.22 0.83
## Year2001 0.05043 0.33587 0.15 0.88
## Year2002 -0.15553 0.18274 -0.85 0.40
## Year2003 0.22089 0.17379 1.27 0.21
## Year2004 0.12996 0.17703 0.73 0.46
## Year2005 0.02707 0.18887 0.14 0.89
```

```

## Year2006      -0.07181    0.20027   -0.36    0.72
## Year2007      0.09605    0.18406    0.52    0.60
## Year2008      0.10460    0.18566    0.56    0.57
## Year2009     -0.00598    0.19445   -0.03    0.98
## Year2010     -0.14401    0.18832   -0.76    0.45
## Year2011     -0.01381    0.17769   -0.08    0.94
## Year2012     -0.03141    0.19162   -0.16    0.87
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.354
## Multiple R-squared:  0.0757, Adjusted R-squared:  -0.0114
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 190 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.320  0.869  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      4.76e-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 1.297 1      1.139
## Year              1.297 16      1.008

```

Residuals from first author



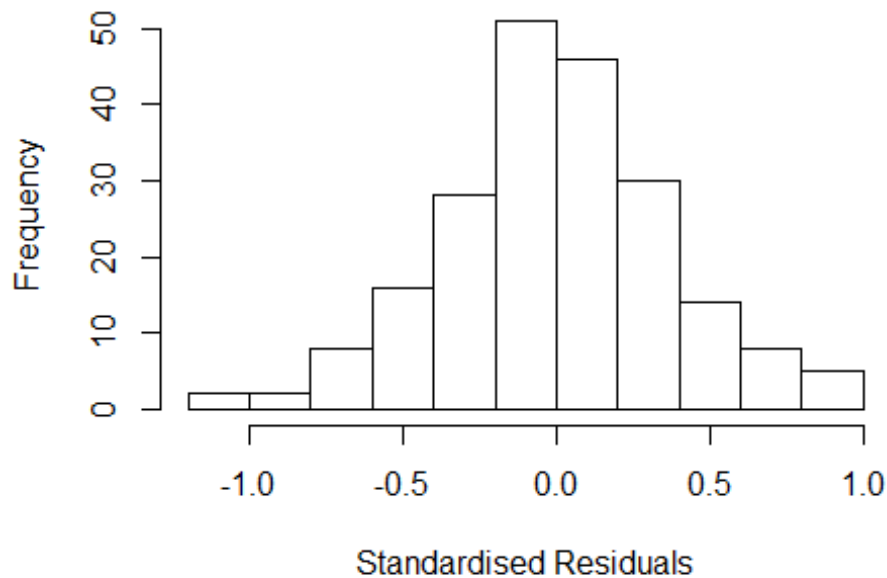
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.09588 -0.22688 -0.00968 0.23118 0.96165
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.090245 0.163550 6.67 2.7e-10 ***
## FirstAuthorFemale1 -0.023259 0.060583 -0.38 0.70
## Year1997 -0.248712 0.248694 -1.00 0.32
## Year1998 -0.000345 0.232442 0.00 1.00
## Year1999 0.028894 0.272060 0.11 0.92
## Year2000 0.051106 0.231712 0.22 0.83
## Year2001 0.050074 0.337434 0.15 0.88
## Year2002 -0.153530 0.182730 -0.84 0.40
## Year2003 0.221721 0.174115 1.27 0.20
## Year2004 0.132624 0.176546 0.75 0.45
## Year2005 0.028775 0.188717 0.15 0.88
## Year2006 -0.068651 0.198414 -0.35 0.73
```

```

## Year2007          0.096820    0.184805    0.52    0.60
## Year2008          0.105805    0.186192    0.57    0.57
## Year2009         -0.003750    0.192433   -0.02    0.98
## Year2010         -0.140574    0.185621   -0.76    0.45
## Year2011         -0.012589    0.177726   -0.07    0.94
## Year2012         -0.027396    0.188122   -0.15    0.88
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.352
## Multiple R-squared:  0.0758, Adjusted R-squared:  -0.00608
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 192 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.312  0.867  0.955  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      4.76e-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 1.45 1      1.204
## Year              1.45 16      1.012

```

Residuals from last author



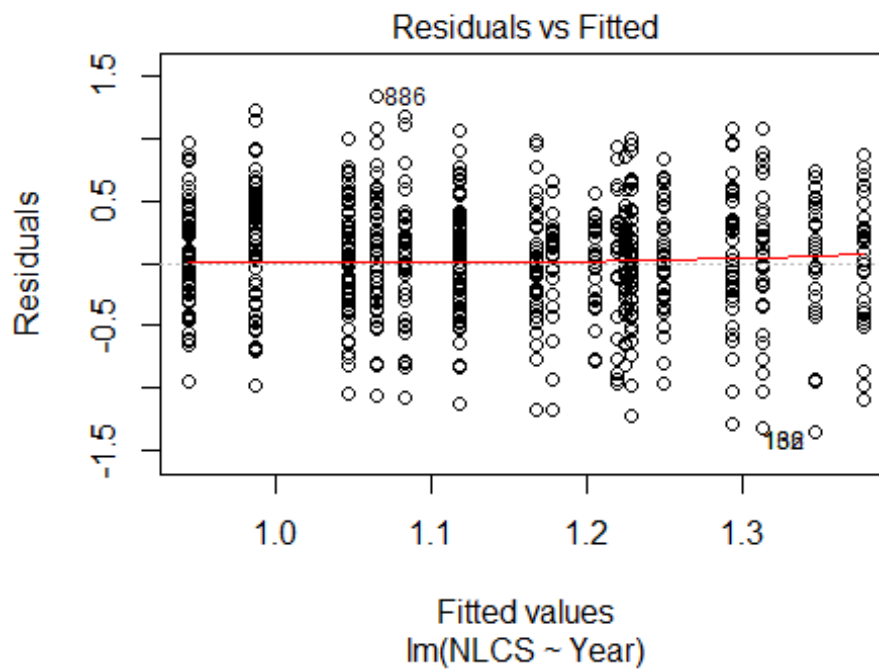
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.10860 -0.23421 -0.00981  0.23693  0.96627
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.08669    0.16706   6.50 6.6e-10 ***
## LastAuthorFemale1  0.00319    0.05562   0.06  0.95
## Year1997       -0.25445    0.25153  -1.01  0.31
## Year1998       -0.00633    0.23615  -0.03  0.98
## Year1999        0.02192    0.27354   0.08  0.94
## Year2000        0.05069    0.23281   0.22  0.83
## Year2001        0.03768    0.34457   0.11  0.91
## Year2002       -0.15329    0.18595  -0.82  0.41
## Year2003        0.22069    0.17691   1.25  0.21
## Year2004        0.13031    0.18064   0.72  0.47
## Year2005        0.02809    0.19198   0.15  0.88
## Year2006       -0.06974    0.20299  -0.34  0.73
```

```

## Year2007          0.08582      0.18594      0.46      0.64
## Year2008          0.09750      0.18831      0.52      0.61
## Year2009         -0.01156      0.19767     -0.06      0.95
## Year2010         -0.14895      0.19004     -0.78      0.43
## Year2011         -0.02179      0.17917     -0.12      0.90
## Year2012         -0.03566      0.19563     -0.18      0.86
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.355
## Multiple R-squared:  0.0753, Adjusted R-squared:  -0.00661
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 192 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.309  0.866  0.953  0.900  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      4.76e-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: 210"
## [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
##   50   53   46   39   54   46   59   47   37   55   61   72   72   57   80
## 2011 2012
##   86   87
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   38   40   28   26   30   20   44   41   33   48   55   66   57   47   65
## 2011 2012

```

```
##      77      76
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   29   36   21   18   23   20   36   30   32   41   45   58   50   40   62
## 2011 2012
##   67   67
## [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
```

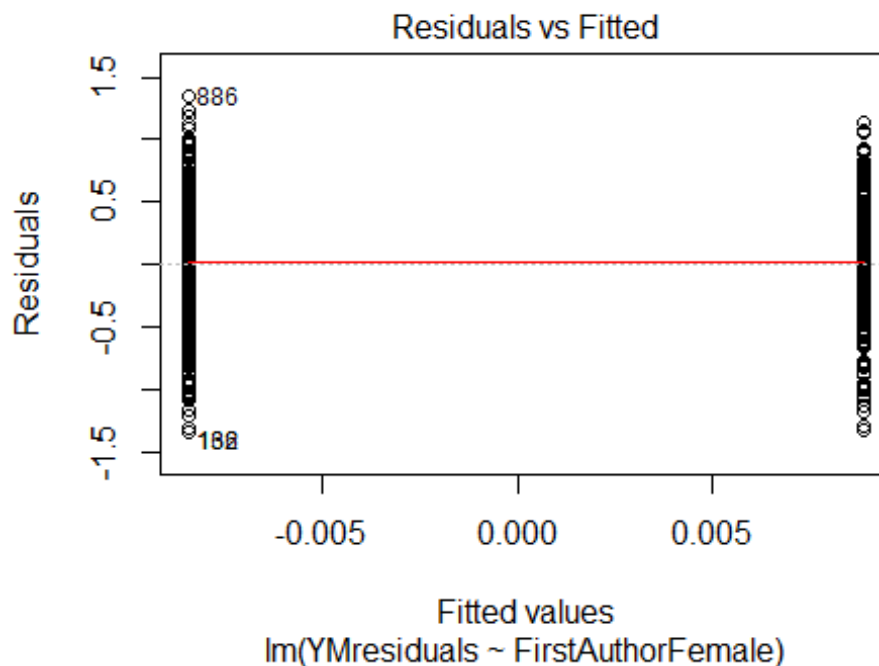


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.9, df = 1, p-value = 0.002
## [1] "Female first author team size 2018 geometric mean: 2.79899807334316"
## [1] "Male first author team size 2018 geometric mean: 2.86855121814652"
## 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 = 450, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.53938929667164"
## [1] "Male last author team size 2018 geometric mean: 3.06684908643138"

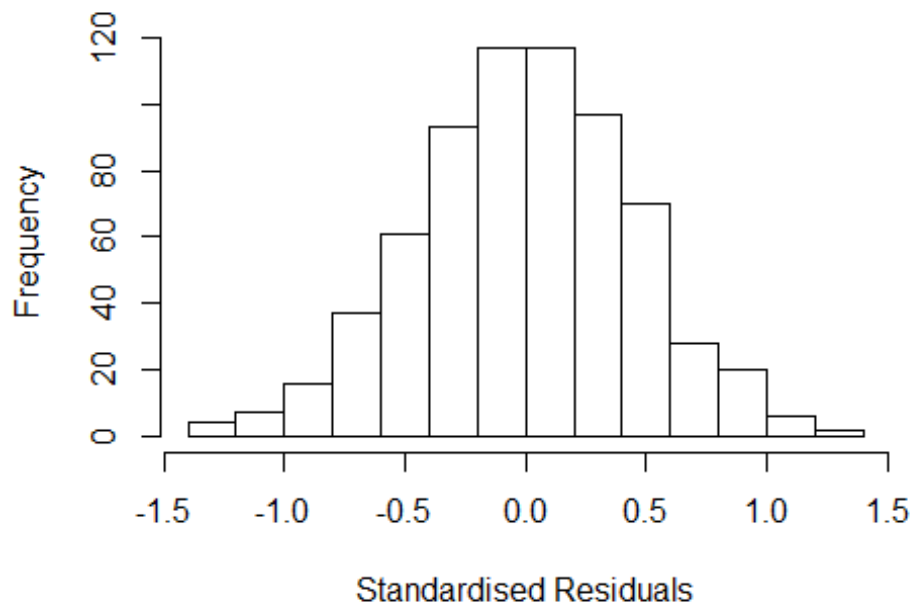
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 420, p-value = 0.2
## 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	1.151	1	1.073
LastAuthorFemale	1.253	1	1.119
UniqueAuthors	1.422	4	1.045
Year	1.837	16	1.019

Residuals from first and last author and team size



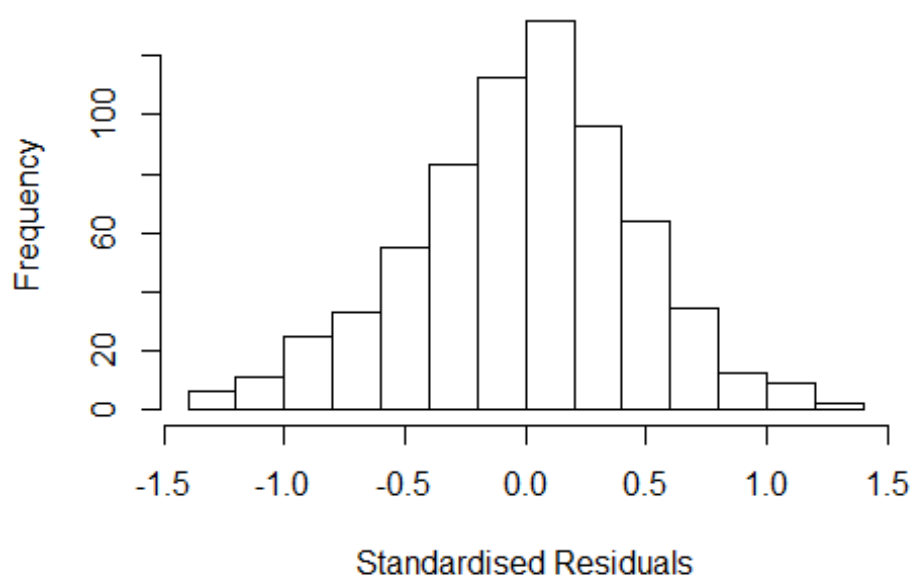
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.32614 -0.30431 0.00229 0.30057 1.24744
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9719 0.1163 8.35 4.0e-16 ***
## FirstAuthorFemale1 0.0848 0.0368 2.30 0.02153 *
## LastAuthorFemale1 0.0378 0.0379 1.00 0.31948
## UniqueAuthors2 0.2165 0.0591 3.66 0.00027 ***
## UniqueAuthors3 0.2760 0.0595 4.64 4.3e-06 ***
## UniqueAuthors4 0.3467 0.0641 5.41 8.7e-08 ***
## UniqueAuthors5 0.3745 0.0611 6.13 1.5e-09 ***
## Year1997 0.0380 0.1425 0.27 0.78991
## Year1998 0.1000 0.1757 0.57 0.56961
## Year1999 -0.0262 0.1463 -0.18 0.85784
```

```

## Year2000          0.0443      0.1467      0.30  0.76269
## Year2001          0.0291      0.1870      0.16  0.87625
## Year2002         -0.0894      0.1399     -0.64  0.52288
## Year2003         -0.0908      0.1319     -0.69  0.49161
## Year2004         -0.0389      0.1237     -0.31  0.75300
## Year2005         -0.1060      0.1292     -0.82  0.41230
## Year2006         -0.0591      0.1170     -0.50  0.61381
## Year2007         -0.2547      0.1261     -2.02  0.04384 *
## Year2008         -0.2673      0.1274     -2.10  0.03632 *
## Year2009         -0.2241      0.1285     -1.74  0.08156 .
## Year2010         -0.3126      0.1354     -2.31  0.02129 *
## Year2011         -0.1530      0.1281     -1.19  0.23283
## Year2012         -0.3345      0.1221     -2.74  0.00631 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.139, Adjusted R-squared:  0.11
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 75 weights are ~= 1. The remaining 600 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.349  0.866  0.948  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      1.48e-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 1.105 1      1.051
## LastAuthorFemale  1.202 1      1.096
## Year              1.313 16      1.009

```

Residuals from first and last author



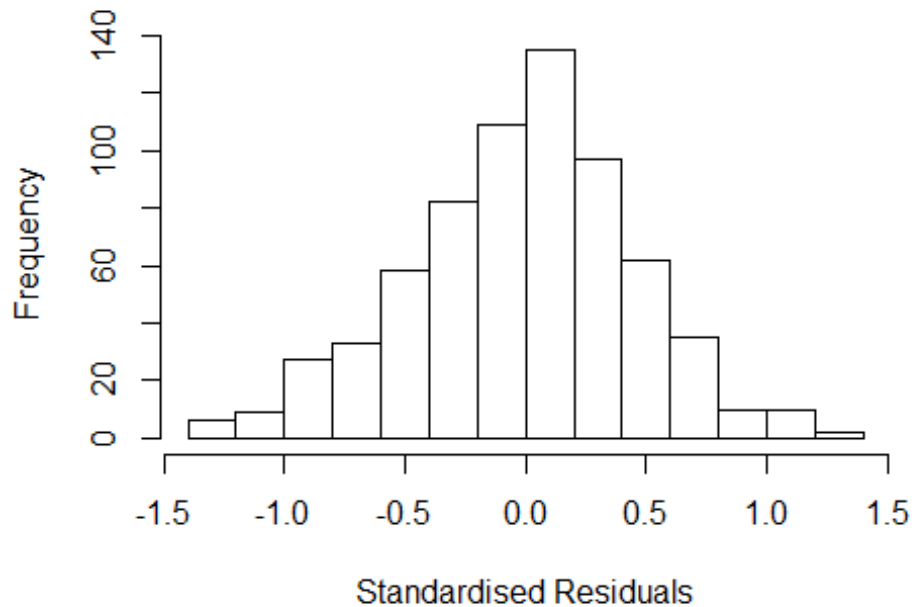
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3669 -0.3037  0.0203  0.2913  1.2933
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.17653    0.12046   9.77   <2e-16 ***
## FirstAuthorFemale1 0.10233    0.03767   2.72   0.0068 **
## LastAuthorFemale1 0.02627    0.03875   0.68   0.4980
## Year1997         0.06176    0.15375   0.40   0.6880
## Year1998         0.11859    0.17961   0.66   0.5093
## Year1999        -0.02976    0.15922  -0.19   0.8518
## Year2000        -0.01187    0.16105  -0.07   0.9413
## Year2001         0.01504    0.19502   0.08   0.9385
## Year2002        -0.05477    0.14808  -0.37   0.7116
## Year2003        -0.04454    0.14228  -0.31   0.7544
## Year2004        -0.01667    0.13223  -0.13   0.8997
## Year2005        -0.07844    0.13868  -0.57   0.5718
```

```

## Year2006      -0.00517    0.12579   -0.04    0.9672
## Year2007      -0.22106    0.13508   -1.64    0.1022
## Year2008      -0.22198    0.13620   -1.63    0.1036
## Year2009      -0.20187    0.14350   -1.41    0.1600
## Year2010      -0.30342    0.14537   -2.09    0.0373 *
## Year2011      -0.17930    0.13646   -1.31    0.1893
## Year2012      -0.29754    0.13299   -2.24    0.0256 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.0736, Adjusted R-squared:  0.0482
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 67 weights are ~= 1. The remaining 608 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.326  0.859   0.949   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      1.48e-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

```

Residuals from first author



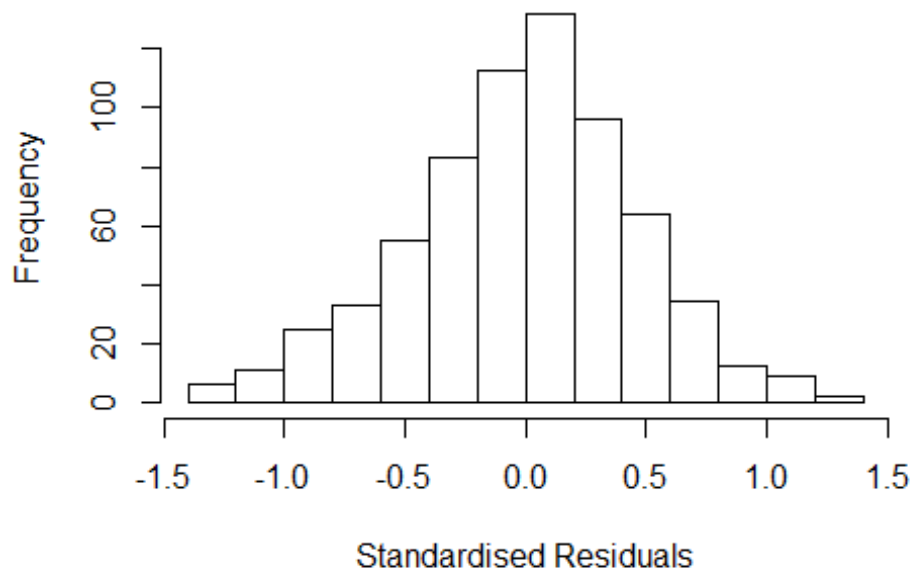
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.348 -0.309  0.018  0.288  1.285
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.18197    0.12011     9.84  <2e-16 ***
## FirstAuthorFemale1 0.10726    0.03789     2.83  0.0048 **
## Year1997         0.05880    0.15344     0.38  0.7017
## Year1998         0.11854    0.17926     0.66  0.5087
## Year1999        -0.03490    0.15870    -0.22  0.8260
## Year2000        -0.00447    0.16060    -0.03  0.9778
## Year2001         0.01904    0.19448     0.10  0.9221
## Year2002        -0.04801    0.14781    -0.32  0.7454
## Year2003        -0.03873    0.14247    -0.27  0.7858
## Year2004        -0.01285    0.13214    -0.10  0.9225
## Year2005        -0.07542    0.13935    -0.54  0.5885
## Year2006        -0.00272    0.12579    -0.02  0.9827
```

```

## Year2007          -0.21800    0.13515   -1.61    0.1072
## Year2008          -0.21995    0.13608   -1.62    0.1065
## Year2009          -0.19849    0.14347   -1.38    0.1670
## Year2010          -0.29719    0.14469   -2.05    0.0404 *
## Year2011          -0.17392    0.13621   -1.28    0.2021
## Year2012          -0.29316    0.13261   -2.21    0.0274 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.0723, Adjusted R-squared:  0.0483
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 67 weights are ~= 1. The remaining 608 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.338  0.859  0.950  0.897  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      1.48e-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 1.182 1      1.087
## Year              1.182 16      1.005

```

Residuals from last author



```
## [1] "List of 0 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
## -1.3709 -0.3091 0.0321 0.2897 1.2500
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.19317 0.12277 9.72 <2e-16 ***
## LastAuthorFemale1 0.04549 0.03904 1.17 0.244
## Year1997 0.09246 0.15488 0.60 0.551
## Year1998 0.13221 0.17485 0.76 0.450
## Year1999 0.02069 0.15669 0.13 0.895
## Year2000 0.04419 0.16088 0.27 0.784
## Year2001 0.04380 0.19598 0.22 0.823
## Year2002 -0.03109 0.14913 -0.21 0.835
## Year2003 -0.02278 0.14267 -0.16 0.873
## Year2004 0.00553 0.13434 0.04 0.967
## Year2005 -0.05488 0.13928 -0.39 0.694
## Year2006 0.01153 0.12788 0.09 0.928
```

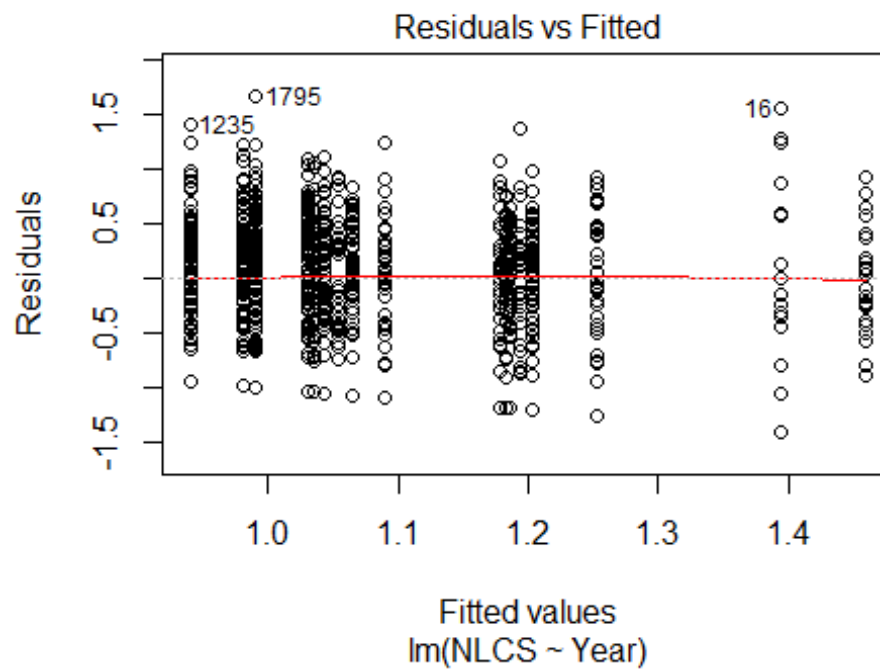


```

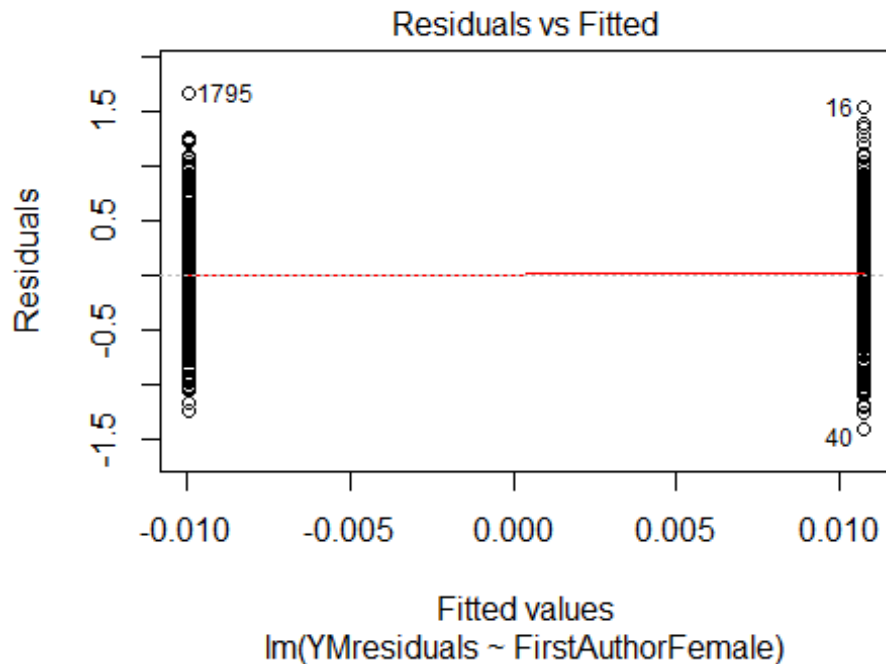
## Year2007          -0.19884      0.13688    -1.45      0.147
## Year2008          -0.19072      0.13919    -1.37      0.171
## Year2009          -0.17518      0.14613    -1.20      0.231
## Year2010          -0.26558      0.14736    -1.80      0.072 .
## Year2011          -0.14318      0.13697    -1.05      0.296
## Year2012          -0.25645      0.13419    -1.91      0.056 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.0641, Adjusted R-squared:  0.0399
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 63 weights are ~= 1. The remaining 612 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.324  0.859  0.948  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      1.48e-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: 675"
## [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
##   31   50   54   59   54   66   57   54   54   75   77   73   99  145  153
## 2011 2012
##  170  166
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   18   29   33   36   30   28   39   36   42   59   57   62   75  102  116
## 2011 2012

```

```
## 145 137
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 16 27 32 34 27 25 34 27 42 48 48 49 67 89 105
## 2011 2012
## 129 124
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 24, df = 16, p-value = 0.08
```

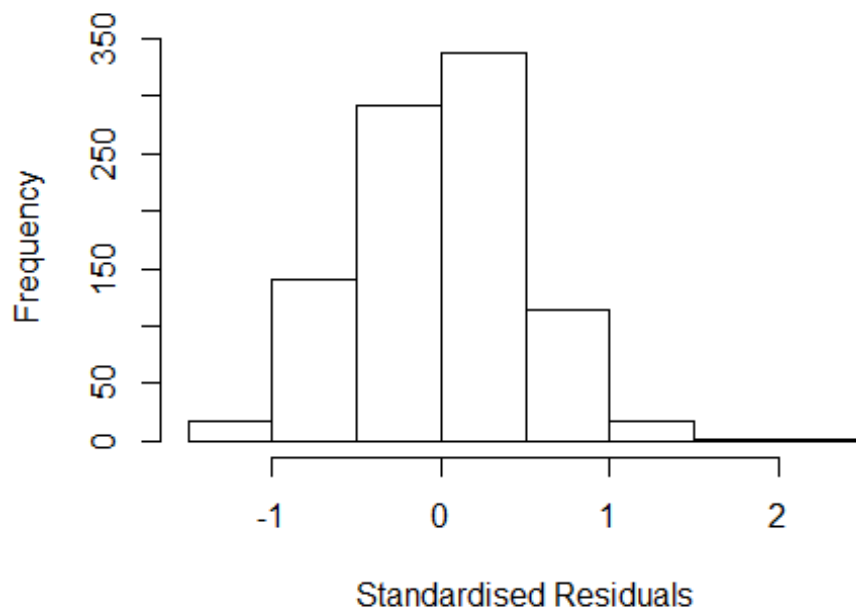


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.62, df = 1, p-value = 0.4
```



```
## [1] "Female first author team size 2018 geometric mean: 3.7479788081126"
## [1] "Male first author team size 2018 geometric mean: 3.89355853816901"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.6935081885541"
## [1] "Male last author team size 2018 geometric mean: 3.91228481669449"
##
## 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 1.141 1      1.068
## LastAuthorFemale  1.211 1      1.100
## UniqueAuthors    1.411 4      1.044
## Year              1.677 16     1.016
```

Residuals from first and last author and team size



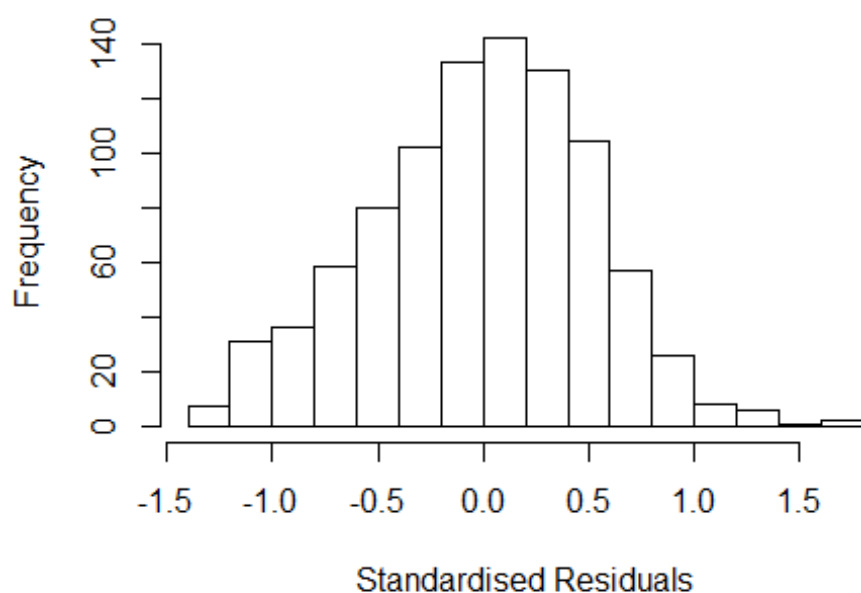
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3906 -0.3363 0.0198 0.3249 2.0624
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.88857 0.23485 3.78 0.00016 ***
## FirstAuthorFemale1 -0.04261 0.03413 -1.25 0.21225
## LastAuthorFemale1 -0.01252 0.03541 -0.35 0.72378
## UniqueAuthors2 0.41645 0.05762 7.23 1.0e-12 ***
## UniqueAuthors3 0.45040 0.05522 8.16 1.2e-15 ***
## UniqueAuthors4 0.51202 0.05904 8.67 < 2e-16 ***
## UniqueAuthors5 0.50199 0.05328 9.42 < 2e-16 ***
## Year1997 -0.01621 0.25102 -0.06 0.94851
## Year1998 0.08308 0.25529 0.33 0.74492
## Year1999 -0.02603 0.24305 -0.11 0.91474
```

```

## Year2000      -0.17501    0.24813   -0.71  0.48079
## Year2001      0.25514    0.24789    1.03  0.30363
## Year2002     -0.10623    0.24698   -0.43  0.66720
## Year2003     -0.03558    0.25001   -0.14  0.88685
## Year2004     -0.00289    0.24299   -0.01  0.99053
## Year2005     -0.22816    0.23913   -0.95  0.34026
## Year2006     -0.06763    0.23980   -0.28  0.77799
## Year2007     -0.20526    0.24584   -0.83  0.40398
## Year2008     -0.22597    0.23886   -0.95  0.34438
## Year2009     -0.26271    0.23539   -1.12  0.26468
## Year2010     -0.24236    0.23567   -1.03  0.30405
## Year2011     -0.23124    0.23507   -0.98  0.32551
## Year2012     -0.25499    0.23604   -1.08  0.28031
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.496
## Multiple R-squared:  0.178, Adjusted R-squared:  0.158
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 77 weights are ~= 1. The remaining 846 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0449 0.8750 0.9480 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      1.08e-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 1.126 1      1.061
## LastAuthorFemale  1.144 1      1.070
## Year              1.187 16      1.005

```

Residuals from first and last author



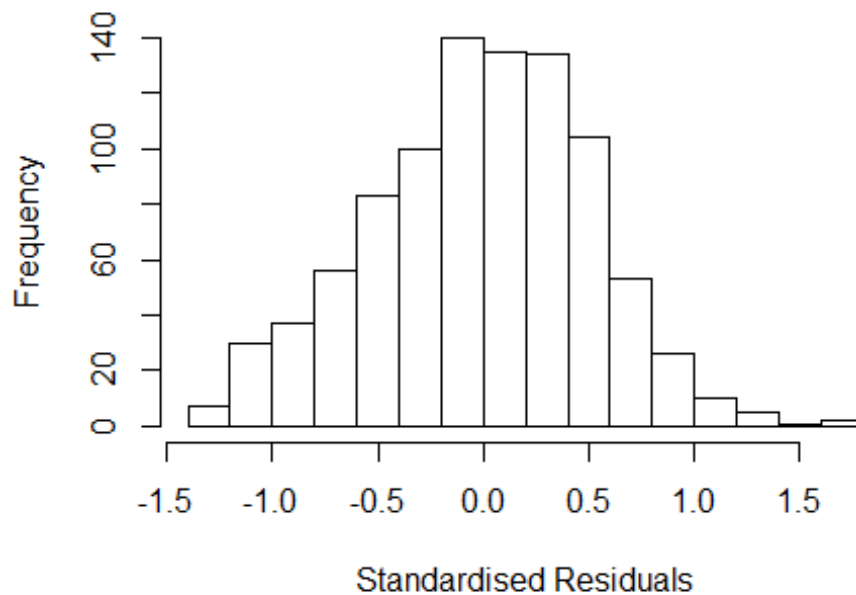
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3175 -0.3627 0.0163 0.3507 1.7194
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.23164 0.27026 4.56 5.9e-06 ***
## FirstAuthorFemale1 -0.01758 0.03636 -0.48 0.63
## LastAuthorFemale1 -0.03651 0.03722 -0.98 0.33
## Year1997 -0.02887 0.28736 -0.10 0.92
## Year1998 0.08587 0.29583 0.29 0.77
## Year1999 -0.04074 0.28270 -0.14 0.89
## Year2000 -0.19265 0.29205 -0.66 0.51
## Year2001 0.24421 0.28057 0.87 0.38
## Year2002 -0.10310 0.28617 -0.36 0.72
## Year2003 -0.01780 0.28766 -0.06 0.95
## Year2004 -0.00413 0.27794 -0.01 0.99
## Year2005 -0.20266 0.27571 -0.74 0.46
```

```

## Year2006          0.01050      0.27557      0.04      0.97
## Year2007          -0.12661      0.27863     -0.45      0.65
## Year2008          -0.17144      0.27557     -0.62      0.53
## Year2009          -0.25986      0.27366     -0.95      0.34
## Year2010          -0.17275      0.27372     -0.63      0.53
## Year2011          -0.21173      0.27204     -0.78      0.44
## Year2012          -0.20410      0.27378     -0.75      0.46
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0479, Adjusted R-squared:  0.0289
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 77 weights are ~= 1. The remaining 846 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.253  0.871  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      1.08e-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 1.081 1          1.040
## Year              1.081 16          1.002

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3068 -0.3531 0.0179 0.3438 1.7230
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22805 0.26945 4.56 5.9e-06 ***
## FirstAuthorFemale1 -0.02745 0.03580 -0.77 0.44
## Year1997 -0.03819 0.28584 -0.13 0.89
## Year1998 0.07880 0.29428 0.27 0.79
## Year1999 -0.04461 0.28196 -0.16 0.87
## Year2000 -0.19999 0.29061 -0.69 0.49
## Year2001 0.23850 0.27933 0.85 0.39
## Year2002 -0.10714 0.28468 -0.38 0.71
## Year2003 -0.01424 0.28691 -0.05 0.96
## Year2004 -0.00836 0.27659 -0.03 0.98
## Year2005 -0.20672 0.27459 -0.75 0.45
## Year2006 0.00771 0.27451 0.03 0.98
```

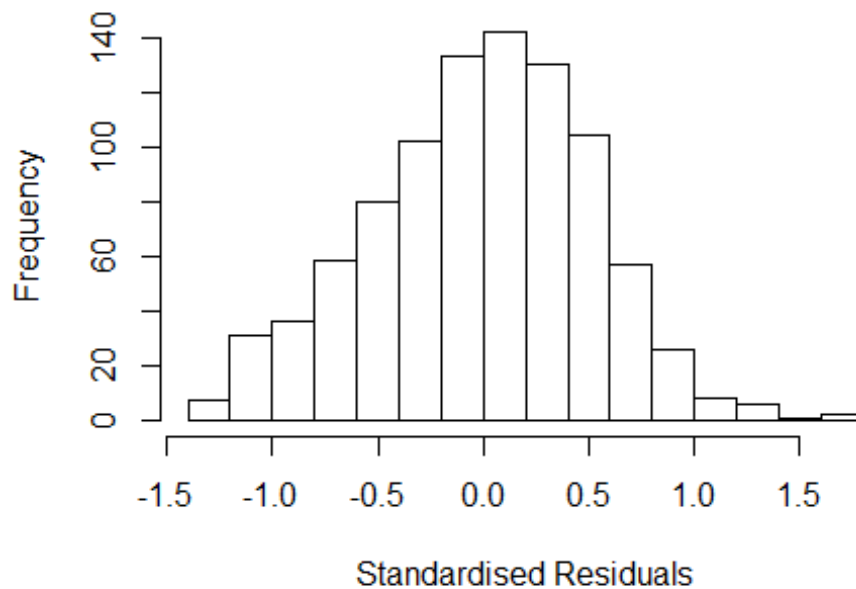


```

## Year2007      -0.13393    0.27665   -0.48    0.63
## Year2008      -0.17532    0.27451   -0.64    0.52
## Year2009      -0.26748    0.27232   -0.98    0.33
## Year2010      -0.18079    0.27202   -0.66    0.51
## Year2011      -0.22146    0.27060   -0.82    0.41
## Year2012      -0.21483    0.27207   -0.79    0.43
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.515
## Multiple R-squared:  0.0471, Adjusted R-squared:  0.0292
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 74 weights are ~= 1. The remaining 849 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.241  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      1.08e-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 1.102 1          1.050
## Year            1.102 16          1.003

```

Residuals from last author



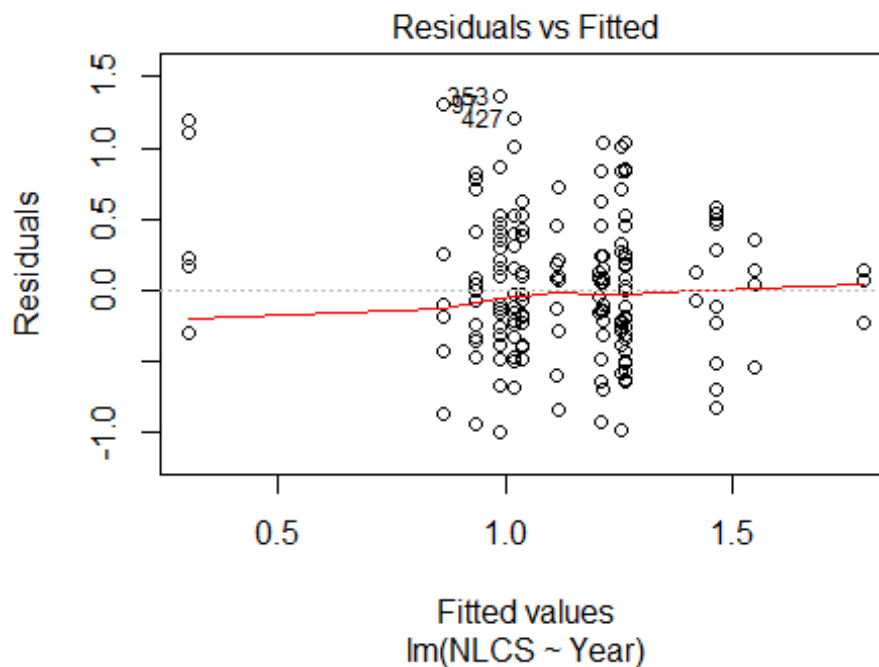
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3100 -0.3581 0.0178 0.3487 1.7295
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.22155 0.26949 4.53 6.6e-06 ***
## LastAuthorFemale1 -0.04125 0.03668 -1.12 0.26
## Year1997 -0.02385 0.28846 -0.08 0.93
## Year1998 0.08844 0.29743 0.30 0.77
## Year1999 -0.03561 0.28342 -0.13 0.90
## Year2000 -0.18810 0.29318 -0.64 0.52
## Year2001 0.24814 0.28207 0.88 0.38
## Year2002 -0.09698 0.28681 -0.34 0.74
## Year2003 -0.01399 0.28918 -0.05 0.96
## Year2004 0.00162 0.27873 0.01 1.00
## Year2005 -0.20158 0.27781 -0.73 0.47
## Year2006 0.01416 0.27700 0.05 0.96
```

```

## Year2007          -0.12248      0.27997   -0.44      0.66
## Year2008          -0.16886      0.27726   -0.61      0.54
## Year2009          -0.25800      0.27547   -0.94      0.35
## Year2010          -0.16996      0.27544   -0.62      0.54
## Year2011          -0.20914      0.27375   -0.76      0.45
## Year2012          -0.20068      0.27536   -0.73      0.47
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.514
## Multiple R-squared:  0.0479, Adjusted R-squared:  0.03
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 74 weights are ~= 1. The remaining 849 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.234  0.868  0.949  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      1.08e-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: 923"
## [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
##    6   14   14   22   11   18   12   8   22   18   26   18   24   24   33
## 2011 2012
##   33   21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    6    3   13    3    6    6    5   12   13   16   10   12   14   24
## 2011 2012

```

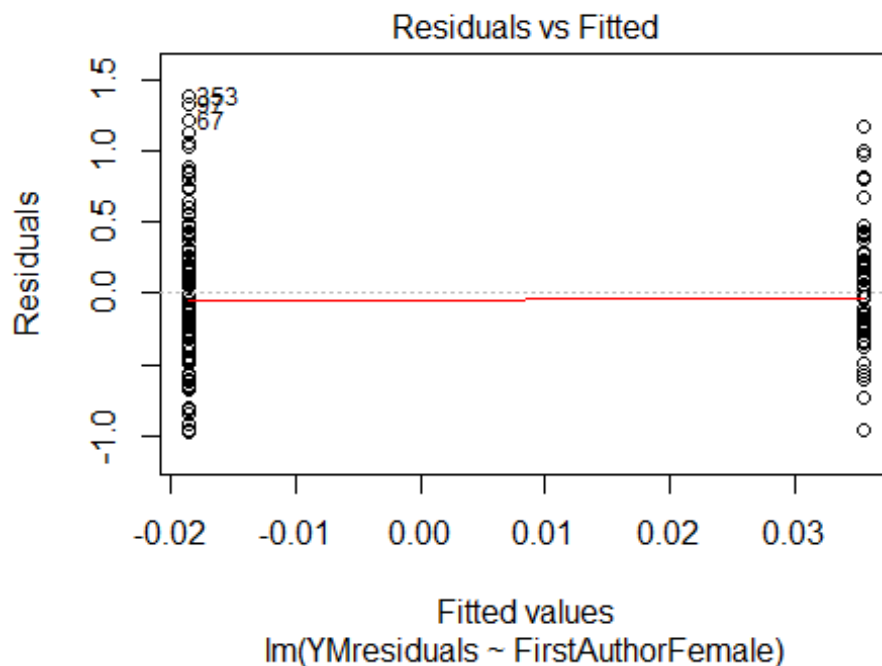
```
## 24 18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 4 3 10 2 4 4 4 7 9 15 10 5 10 16
## 2011 2012
## 21 16
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS 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 = 2.9, df = 1, p-value = 0.09
## [1] "Female first author team size 2018 geometric mean: 5.04537849152229"
## [1] "Male first author team size 2018 geometric mean: 4.39214093036439"
## 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 = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 7.34846922834953"
## [1] "Male last author team size 2018 geometric mean: 4.20662135799604"

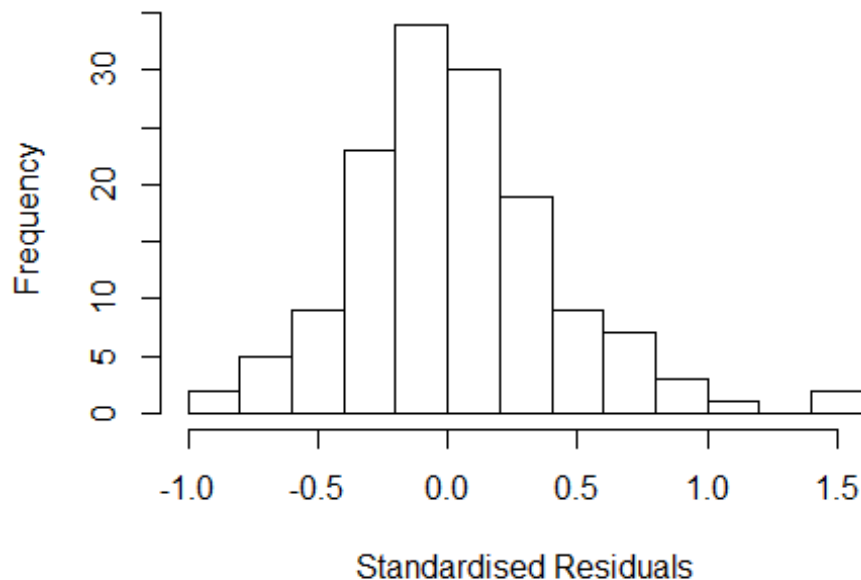
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18, p-value = 0.2
## 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	3.155	1	1.776
LastAuthorFemale	4.819	1	2.195
UniqueAuthors	452.872	4	2.148
Year	1468.010	16	1.256

Residuals from first and last author and team size



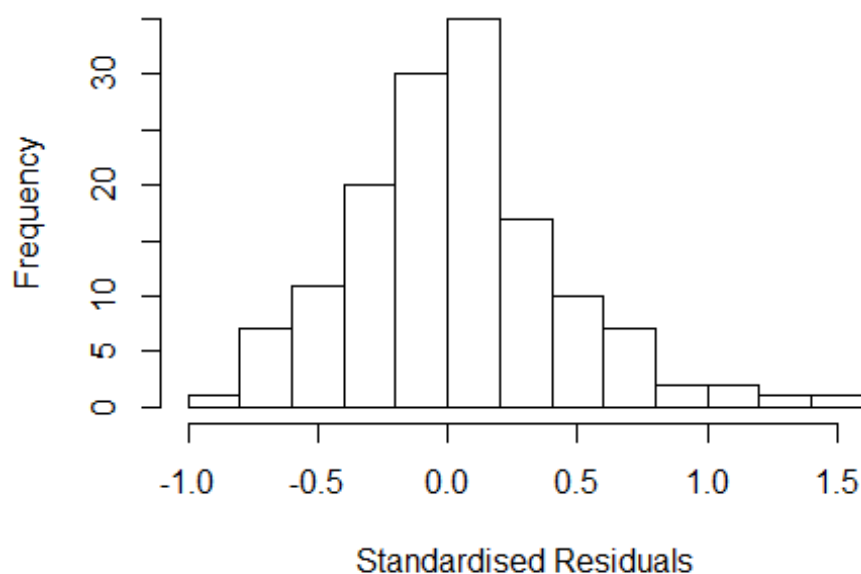
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.831393 -0.230164 -0.000486 0.251757 1.593295
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.47181 0.29346 5.02 1.8e-06 ***
## FirstAuthorFemale1 0.12414 0.08633 1.44 0.15304
## LastAuthorFemale1 0.12492 0.10751 1.16 0.24756
## UniqueAuthors2 -0.08218 0.26236 -0.31 0.75464
## UniqueAuthors3 -0.00302 0.27699 -0.01 0.99131
## UniqueAuthors4 0.02784 0.25400 0.11 0.91289
## UniqueAuthors5 0.17783 0.24863 0.72 0.47585
## Year1997 -0.34041 0.22041 -1.54 0.12509
## Year1998 0.19514 0.18087 1.08 0.28277
## Year1999 -1.65211 0.18490 -8.94 5.5e-15 ***
```

```

## Year2000      -0.23882    0.21420   -1.11  0.26708
## Year2001      -0.86594    0.22581   -3.83  0.00020 ***
## Year2002      -0.28097    0.17058   -1.65  0.10213
## Year2003      -0.58821    0.23688   -2.48  0.01439 *
## Year2004      -0.46872    0.19008   -2.47  0.01507 *
## Year2005      -0.62226    0.23090   -2.69  0.00804 **
## Year2006      -0.61551    0.18913   -3.25  0.00147 **
## Year2007      -0.16939    0.29007   -0.58  0.56034
## Year2008      -0.52732    0.23679   -2.23  0.02780 *
## Year2009      -0.92490    0.24202   -3.82  0.00021 ***
## Year2010      -0.65722    0.21660   -3.03  0.00295 **
## Year2011      -0.37709    0.21033   -1.79  0.07550 .
## Year2012      -0.64790    0.20827   -3.11  0.00233 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.52, Adjusted R-squared:  0.433
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## observation 21 is an outlier with |weight| <= 0.00014 ( < 0.00069);
## 16 weights are ~= 1. The remaining 127 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0147 0.8370 0.9470 0.8810 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           6.94e-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 2.030 1 1.425
## LastAuthorFemale 3.236 1 1.799
## Year 6.457 16 1.060

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -0.8451 -0.2549 0.0194 0.2312 1.4339
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5178 0.1993 7.62 5.6e-12 ***
## FirstAuthorFemale1 0.1254 0.0799 1.57 0.11884
## LastAuthorFemale1 0.1050 0.1040 1.01 0.31446
## Year1997 -0.3960 0.2172 -1.82 0.07064 .
## Year1998 0.1980 0.2378 0.83 0.40668
## Year1999 -1.5388 0.2212 -6.96 1.7e-10 ***
## Year2000 -0.2786 0.2068 -1.35 0.18045
## Year2001 -0.8572 0.2336 -3.67 0.00036 ***
## Year2002 -0.2918 0.2061 -1.42 0.15940
## Year2003 -0.5398 0.2593 -2.08 0.03943 *
## Year2004 -0.4038 0.2255 -1.79 0.07576 .
## Year2005 -0.5197 0.2434 -2.14 0.03470 *
```

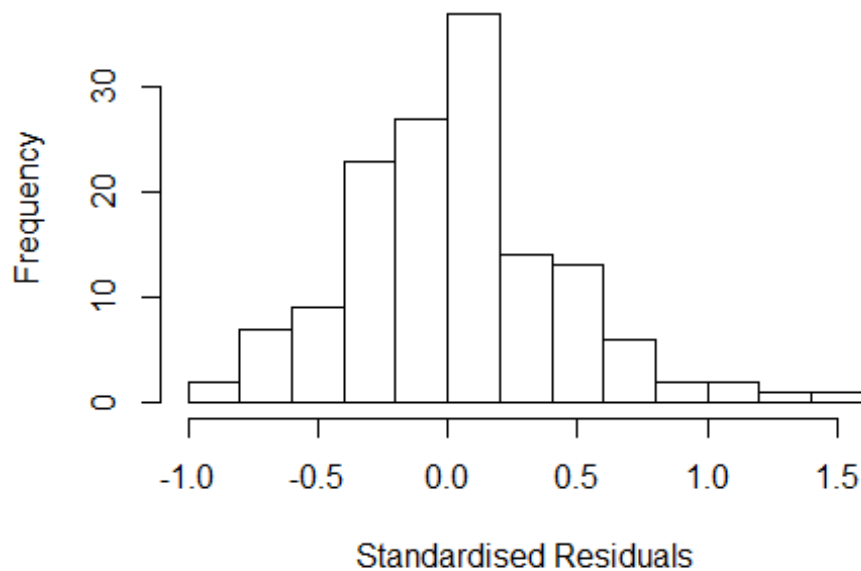


```

## Year2006          -0.5935      0.2194    -2.71  0.00778 **
## Year2007          -0.0899      0.3237    -0.28  0.78162
## Year2008          -0.5331      0.2821    -1.89  0.06109 .
## Year2009          -0.8675      0.2667    -3.25  0.00147 **
## Year2010          -0.6728      0.2333    -2.88  0.00464 **
## Year2011          -0.3132      0.2267    -1.38  0.16961
## Year2012          -0.6566      0.2300    -2.85  0.00504 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.377
## Multiple R-squared:  0.472, Adjusted R-squared:  0.396
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 127 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.115  0.859  0.944  0.889  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      6.94e-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 1.948 1      1.396
## Year      1.948 16      1.021

```

Residuals from first author



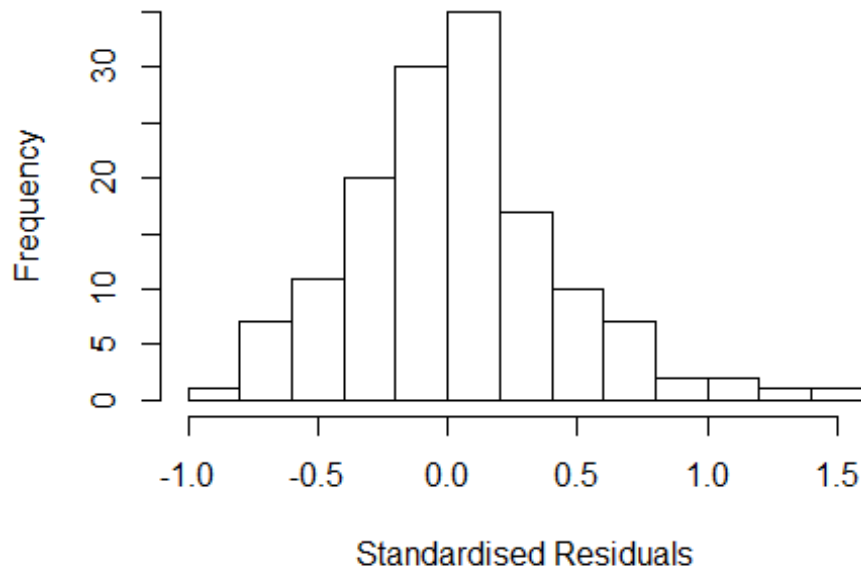
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.8500 -0.2463  0.0167  0.2187  1.4327
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5204     0.1987   7.65 4.5e-12 ***
## FirstAuthorFemale1  0.1213     0.0803   1.51  0.1334
## Year1997         -0.3717     0.2200  -1.69  0.0935 .
## Year1998          0.2344     0.2367   0.99  0.3239
## Year1999         -1.5401     0.2202  -6.99 1.4e-10 ***
## Year2000          -0.2265     0.2038  -1.11  0.2684
## Year2001          -0.8334     0.2240  -3.72  0.0003 ***
## Year2002          -0.2675     0.2005  -1.33  0.1847
## Year2003          -0.5409     0.2588  -2.09  0.0386 *
## Year2004          -0.3880     0.2246  -1.73  0.0865 .
## Year2005          -0.5236     0.2429  -2.16  0.0330 *
## Year2006          -0.5805     0.2189  -2.65  0.0090 **
```

```

## Year2007          -0.0752      0.3331   -0.23   0.8217
## Year2008          -0.4862      0.2630   -1.85   0.0669 .
## Year2009          -0.8703      0.2661   -3.27   0.0014 **
## Year2010          -0.6704      0.2313   -2.90   0.0044 **
## Year2011          -0.2994      0.2285   -1.31   0.1925
## Year2012          -0.6288      0.2281   -2.76   0.0067 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.373
## Multiple R-squared:  0.47,   Adjusted R-squared:  0.398
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 129 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.106  0.853  0.945  0.888  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      6.94e-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 2.387 1          1.545
## Year            2.387 16          1.028

```

Residuals from last author



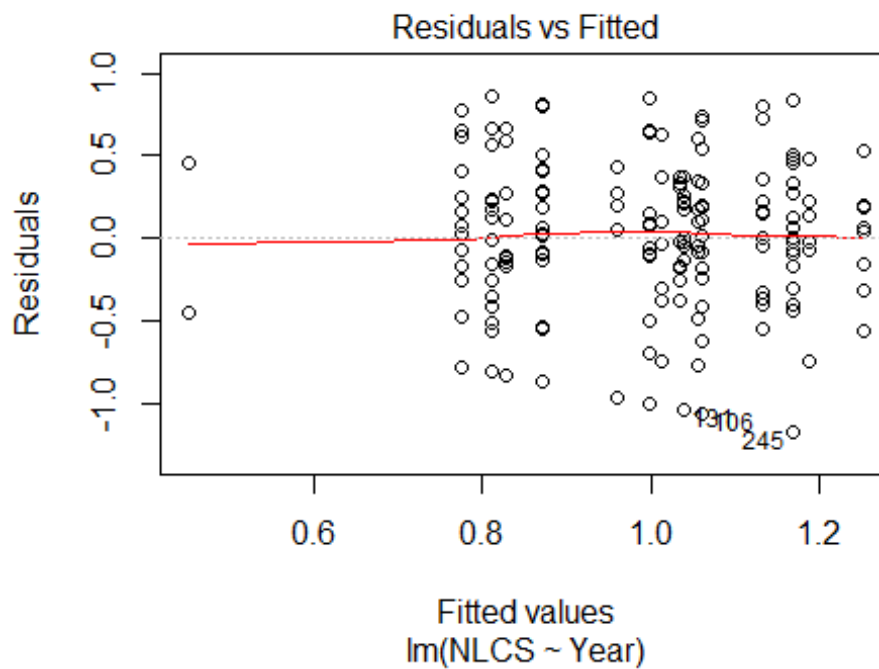
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.86344 -0.20585 -0.00139 0.23703 1.34053
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5734 0.1750 8.99 3.1e-15 ***
## LastAuthorFemale1 0.0966 0.1114 0.87 0.38738
## Year1997 -0.4493 0.1960 -2.29 0.02357 *
## Year1998 0.1853 0.2005 0.92 0.35721
## Year1999 -1.5010 0.1920 -7.82 1.9e-12 ***
## Year2000 -0.2672 0.1864 -1.43 0.15402
## Year2001 -0.8485 0.2269 -3.74 0.00028 ***
## Year2002 -0.3134 0.1844 -1.70 0.09165 .
## Year2003 -0.5575 0.2368 -2.35 0.02008 *
## Year2004 -0.4369 0.2035 -2.15 0.03372 *
## Year2005 -0.5661 0.2314 -2.45 0.01581 *
## Year2006 -0.5921 0.1964 -3.01 0.00311 **
```

```

## Year2007          -0.0739      0.2838   -0.26   0.79510
## Year2008          -0.5609      0.2815   -1.99   0.04844 *
## Year2009          -0.8469      0.2434   -3.48   0.00069 ***
## Year2010          -0.7100      0.2163   -3.28   0.00133 **
## Year2011          -0.3106      0.2116   -1.47   0.14465
## Year2012          -0.6884      0.2175   -3.17   0.00194 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.372
## Multiple R-squared:  0.461, Adjusted R-squared:  0.389
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 122 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.167  0.825   0.939   0.880   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      6.94e-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: 144"
## [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
##   18   18   20   19   22   20   16    9   17   19   27   24   22   18   14
## 2011 2012
##   26   32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    2    8   14   15    9   11    6    8   12   10   18   12    8    6
## 2011 2012

```

```
## 19 22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 1 6 13 13 8 8 4 6 9 10 16 10 8 4
## 2011 2012
## 15 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 = 6, df = 16, p-value = 1
```



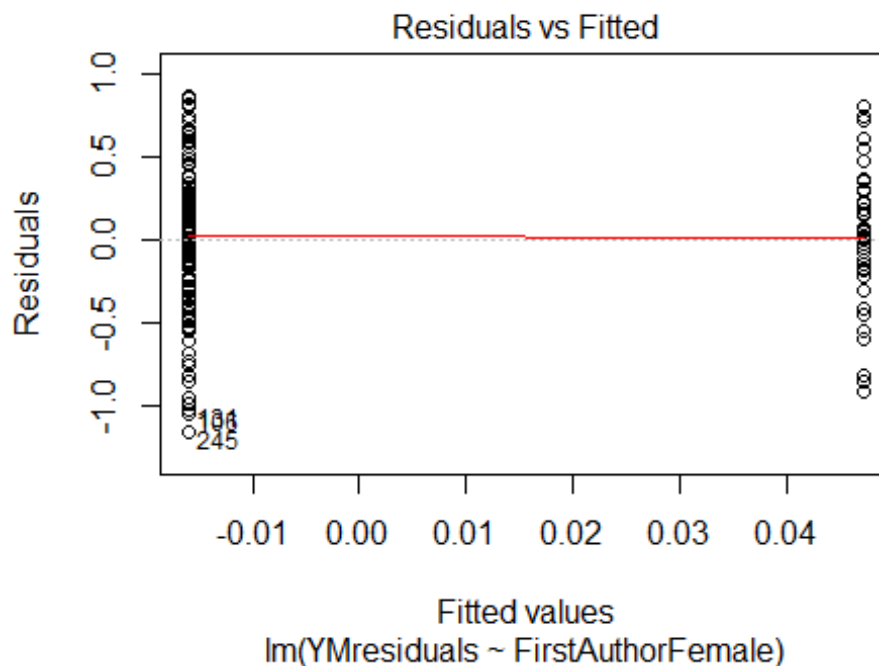
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.28, df = 1, p-value = 0.6

## [1] "Female first author team size 2018 geometric mean: 3.99355148544288"
## [1] "Male first author team size 2018 geometric mean: 2.79816641433953"

## 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 = 57, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.30975091964687"
## [1] "Male last author team size 2018 geometric mean: 3.15879238869588"

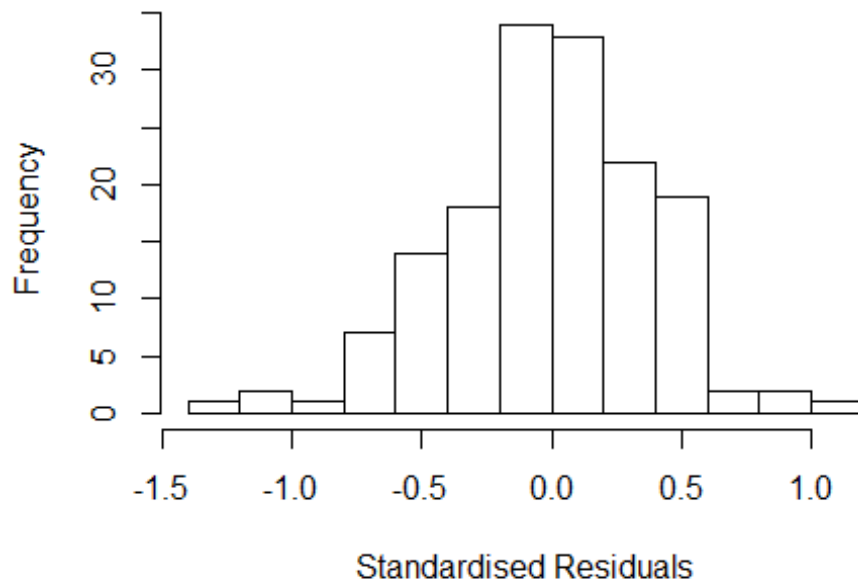
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 29, 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	2.167	1	1.472
LastAuthorFemale	4.889	1	2.211
UniqueAuthors	31.965	4	1.542
Year	68.256	16	1.141

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.2011 -0.2725  0.0133  0.2616  1.0325
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9427    0.2235   4.22 4.5e-05 ***
## FirstAuthorFemale1 -0.0128    0.0795  -0.16  0.87250
## LastAuthorFemale1  0.0307    0.0779   0.39  0.69405
## UniqueAuthors2    0.1768    0.1431   1.24  0.21891
## UniqueAuthors3    0.1883    0.1377   1.37  0.17386
## UniqueAuthors4    0.3336    0.1463   2.28  0.02425 *
## UniqueAuthors5    0.5026    0.1301   3.86  0.00017 ***
## Year1997         -0.2492    0.2640  -0.94  0.34700
## Year1998         -0.0498    0.2321  -0.21  0.83056
## Year1999         -0.1112    0.2638  -0.42  0.67421
```

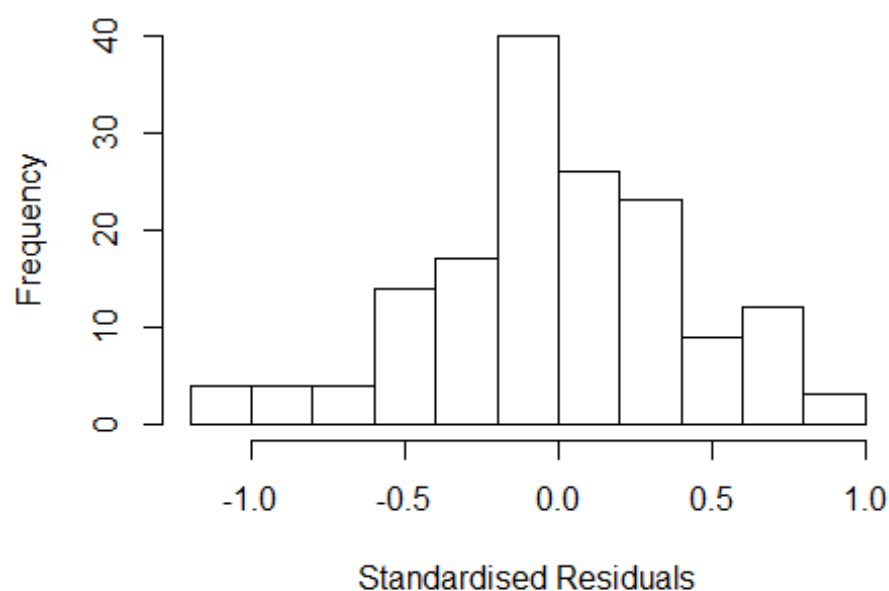


```

## Year2000          -0.1446      0.2378   -0.61  0.54426
## Year2001          -0.1993      0.2465   -0.81  0.42022
## Year2002          -0.2441      0.2307   -1.06  0.29204
## Year2003          -0.3495      0.2822   -1.24  0.21781
## Year2004          -0.0197      0.2796   -0.07  0.94382
## Year2005          -0.4193      0.2472   -1.70  0.09219 .
## Year2006          -0.3905      0.2428   -1.61  0.11013
## Year2007          -0.0661      0.2335   -0.28  0.77770
## Year2008          -0.0579      0.2366   -0.24  0.80697
## Year2009          -0.1859      0.2434   -0.76  0.44653
## Year2010          -0.2512      0.3013   -0.83  0.40591
## Year2011          -0.3962      0.2383   -1.66  0.09866 .
## Year2012          -0.3134      0.2203   -1.42  0.15730
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.403
## Multiple R-squared:  0.244, Adjusted R-squared:  0.119
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 141 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.355  0.883  0.953  0.912  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          6.41e-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 2.426 1 1.558
## LastAuthorFemale 4.823 1 2.196
## Year 9.411 16 1.073

```

Residuals from first and last author



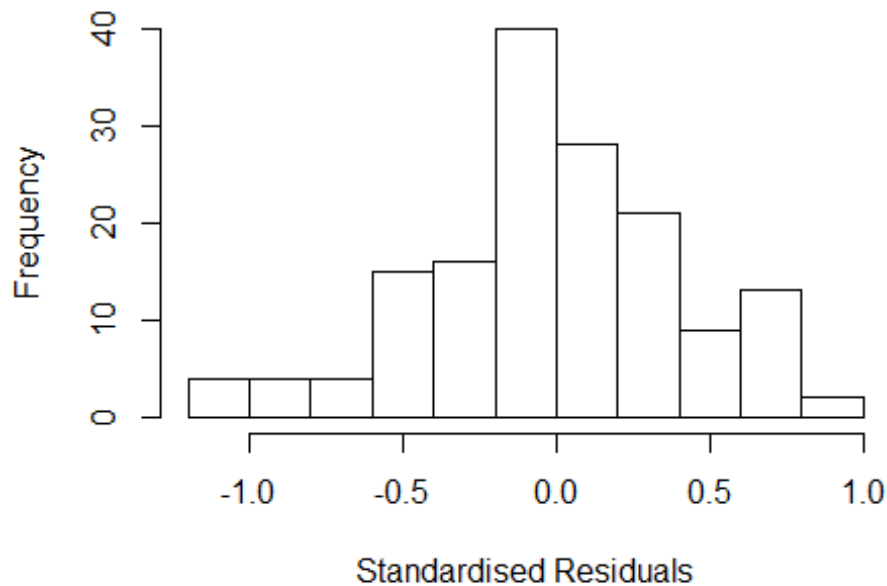
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1879 -0.2482 -0.0134 0.2379 0.8211
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0606 0.1980 5.36 3.5e-07 ***
## FirstAuthorFemale1 0.0162 0.0881 0.18 0.85
## LastAuthorFemale1 0.0227 0.0846 0.27 0.79
## Year1997 -0.1823 0.2229 -0.82 0.41
## Year1998 0.1382 0.2518 0.55 0.58
## Year1999 0.0696 0.2392 0.29 0.77
## Year2000 0.0121 0.2500 0.05 0.96
## Year2001 -0.0380 0.2470 -0.15 0.88
## Year2002 -0.0514 0.2194 -0.23 0.82
## Year2003 -0.2003 0.3563 -0.56 0.57
## Year2004 0.1227 0.2444 0.50 0.62
## Year2005 -0.2356 0.2356 -1.00 0.32
```

```

## Year2006          -0.2413      0.2558    -0.94      0.35
## Year2007           0.1273      0.2193      0.58      0.56
## Year2008           0.0942      0.2421      0.39      0.70
## Year2009          -0.0351      0.2257    -0.16      0.88
## Year2010           0.1009      0.3146      0.32      0.75
## Year2011          -0.2830      0.2296    -1.23      0.22
## Year2012          -0.1815      0.2157    -0.84      0.40
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.416
## Multiple R-squared:  0.108, Adjusted R-squared:  -0.00876
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 141 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.395  0.865   0.964   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      6.41e-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 2.077 1      1.441
## Year              2.077 16      1.023

```

Residuals from first author



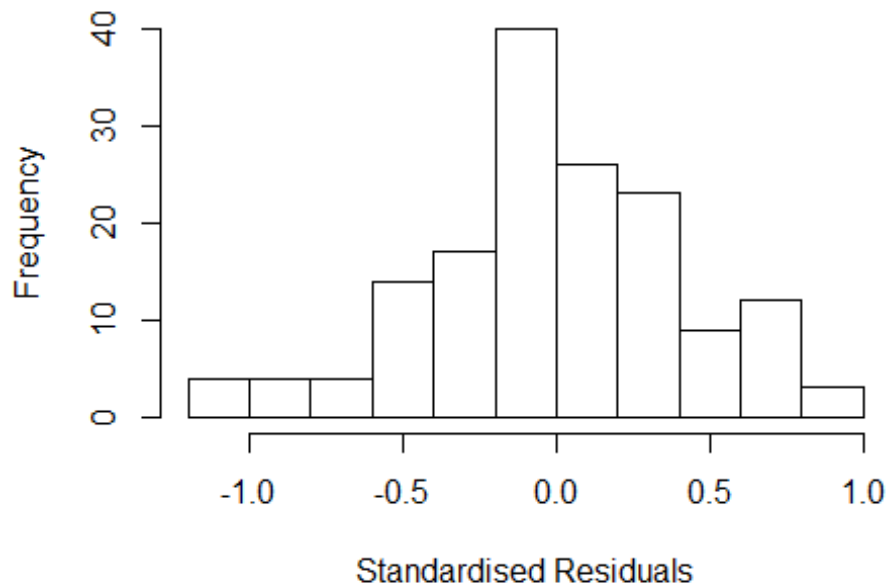
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1883 -0.2483 -0.0114  0.2474  0.8207
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0584     0.1971    5.37 3.3e-07 ***
## FirstAuthorFemale1  0.0211     0.0837    0.25  0.80
## Year1997         -0.1574     0.1971   -0.80  0.43
## Year1998          0.1441     0.2495    0.58  0.56
## Year1999          0.0767     0.2333    0.33  0.74
## Year2000          0.0171     0.2493    0.07  0.95
## Year2001         -0.0330     0.2444   -0.13  0.89
## Year2002         -0.0450     0.2174   -0.21  0.84
## Year2003         -0.1910     0.3435   -0.56  0.58
## Year2004          0.1324     0.2385    0.55  0.58
## Year2005         -0.2292     0.2313   -0.99  0.32
## Year2006         -0.2327     0.2540   -0.92  0.36
```

```

## Year2007          0.1299      0.2180      0.60      0.55
## Year2008          0.0956      0.2412      0.40      0.69
## Year2009         -0.0303      0.2228     -0.14      0.89
## Year2010          0.1011      0.3150      0.32      0.75
## Year2011         -0.2726      0.2212     -1.23      0.22
## Year2012         -0.1789      0.2152     -0.83      0.41
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.417
## Multiple R-squared:  0.108, Adjusted R-squared:  -0.00201
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 139 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.397  0.860  0.962  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      6.41e-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 4.305 1      2.075
## Year            4.305 16      1.047

```

Residuals from last author



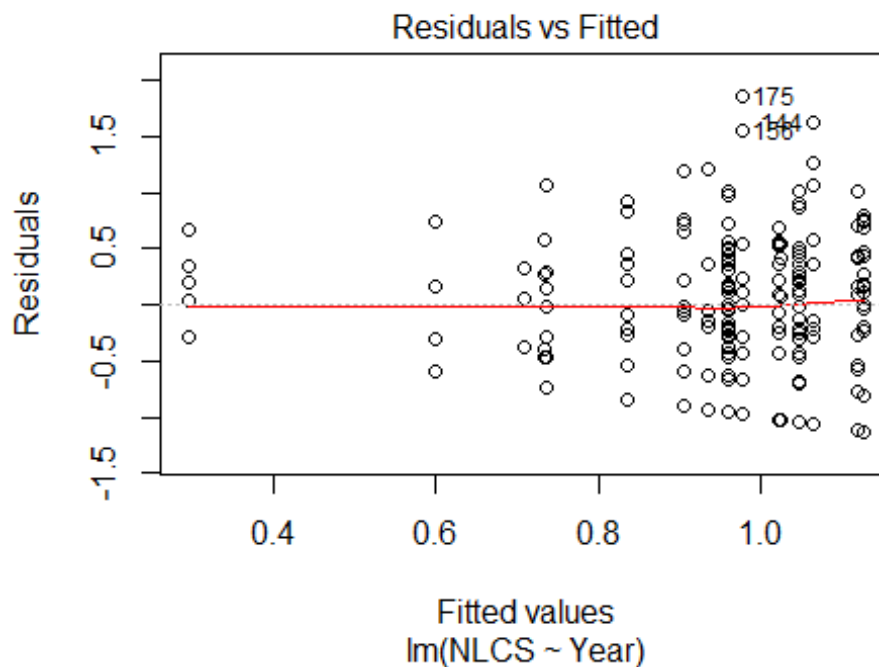
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.19241 -0.25058 -0.00661 0.23841 0.81659
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06730 0.19164 5.57 1.3e-07 ***
## LastAuthorFemale1 0.02733 0.08007 0.34 0.73
## Year1997 -0.19363 0.20946 -0.92 0.36
## Year1998 0.13224 0.24711 0.54 0.59
## Year1999 0.06643 0.23707 0.28 0.78
## Year2000 0.00698 0.24738 0.03 0.98
## Year2001 -0.04343 0.23951 -0.18 0.86
## Year2002 -0.05127 0.21731 -0.24 0.81
## Year2003 -0.20886 0.35184 -0.59 0.55
## Year2004 0.11759 0.24113 0.49 0.63
## Year2005 -0.23687 0.23413 -1.01 0.31
## Year2006 -0.24365 0.25473 -0.96 0.34
```

```

## Year2007      0.12511      0.21680      0.58      0.56
## Year2008      0.09005      0.23860      0.38      0.71
## Year2009     -0.04020      0.22237     -0.18      0.86
## Year2010      0.10046      0.31256      0.32      0.75
## Year2011     -0.28656      0.22759     -1.26      0.21
## Year2012     -0.18075      0.21559     -0.84      0.40
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.109, Adjusted R-squared:  -0.000953
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 141 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.383  0.861  0.963   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      6.41e-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: 156"
## [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
##   13   8   9  11  13  11   5  11  14  18  19  28  15  24  18
## 2011 2012
##   34  35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   3   4  11   8   4   4   8  12  12  12  24  13  18  13
## 2011 2012

```

```
## 24 29
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 3 4 11 7 4 3 8 11 9 10 23 11 15 12
## 2011 2012
## 20 25
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 20, df = 16, p-value = 0.2
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.8, df = 1, p-value = 0.2

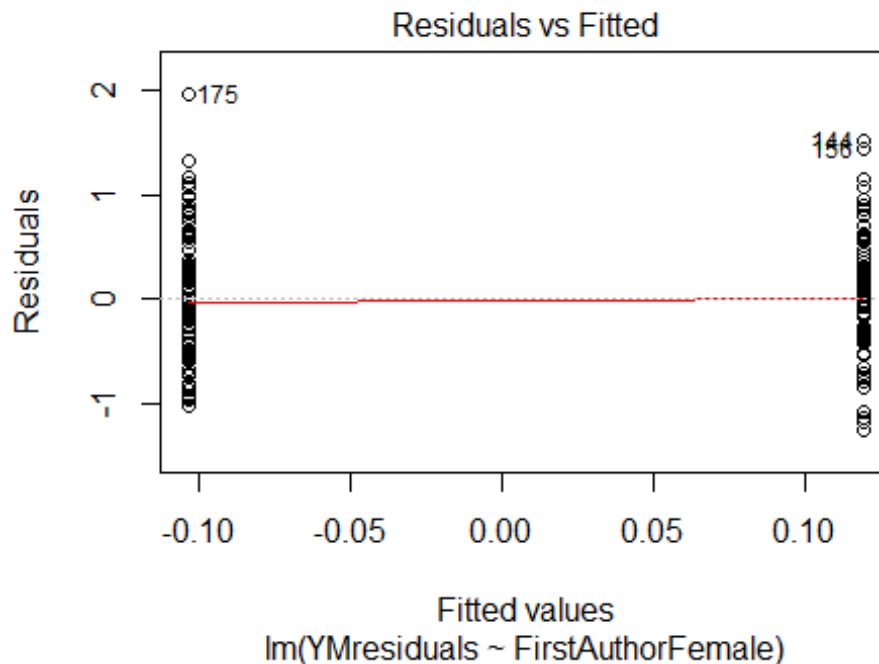
## [1] "Female first author team size 2018 geometric mean: 4.47522342017627"
## [1] "Male first author team size 2018 geometric mean: 5.47722557505166"

## 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 = 12, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male last author team size 2018 geometric mean: 5.48587408073965"

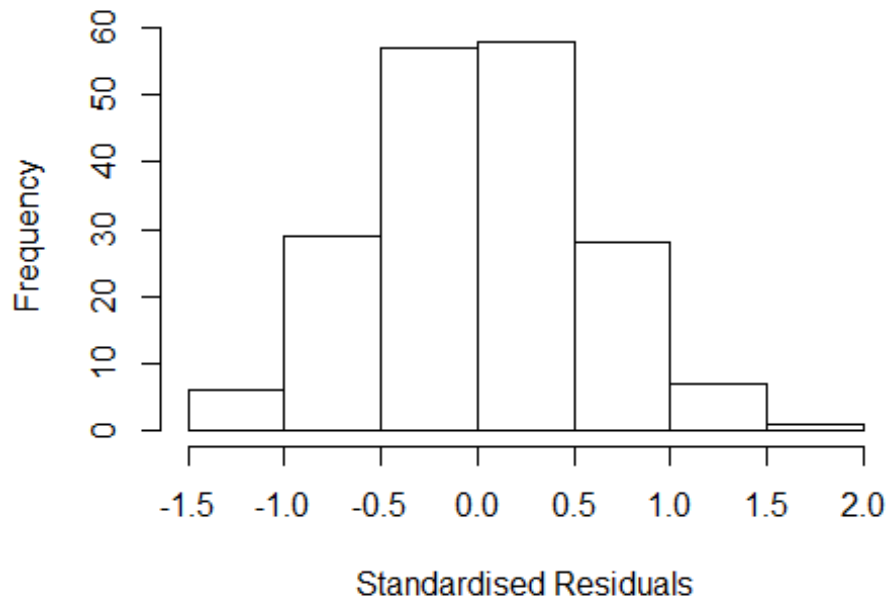
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, 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	1.851	1	1.361
LastAuthorFemale	2.189	1	1.480
UniqueAuthors	6.406	4	1.261
Year	9.928	16	1.074

Residuals from first and last author and team size



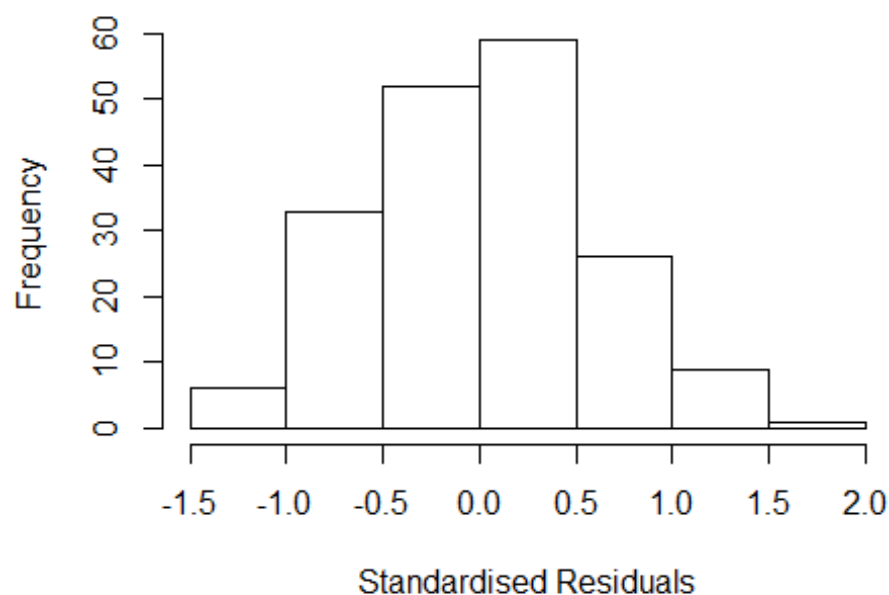
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2849 -0.3933  0.0101  0.3978  1.8831
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.319      0.113    2.81  0.00550 **
## FirstAuthorFemale1 -0.241      0.109   -2.20  0.02923 *
## LastAuthorFemale1 -0.175      0.117   -1.50  0.13563
## UniqueAuthors2      0.103      0.137    0.75  0.45265
## UniqueAuthors3      0.226      0.133    1.71  0.08959 .
## UniqueAuthors4      0.132      0.229    0.58  0.56507
## UniqueAuthors5      0.416      0.135    3.09  0.00235 **
## Year1997           0.552      0.131    4.21  4.2e-05 ***
## Year1998           0.600      0.286    2.09  0.03781 *
## Year1999           0.602      0.208    2.90  0.00430 **
```

```

## Year2000          0.312      0.176      1.77  0.07831 .
## Year2001          0.220      0.325      0.68  0.50053
## Year2002          0.741      0.496      1.50  0.13676
## Year2003          0.685      0.256      2.68  0.00820 **
## Year2004          0.886      0.218      4.07  7.3e-05 ***
## Year2005          0.560      0.558      1.00  0.31666
## Year2006          0.566      0.250      2.26  0.02485 *
## Year2007          0.676      0.175      3.86  0.00016 ***
## Year2008          0.915      0.240      3.81  0.00020 ***
## Year2009          0.774      0.204      3.79  0.00021 ***
## Year2010          0.782      0.252      3.10  0.00225 **
## Year2011          0.697      0.168      4.14  5.5e-05 ***
## Year2012          0.766      0.163      4.71  5.4e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.205, Adjusted R-squared:  0.0971
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 172 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.268  0.881  0.952  0.910  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      5.38e-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 2: First author gender, Last author gender, Year as
factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.754 1      1.325
## LastAuthorFemale  1.606 1      1.267
## Year              1.983 16      1.022

```

Residuals from first and last author



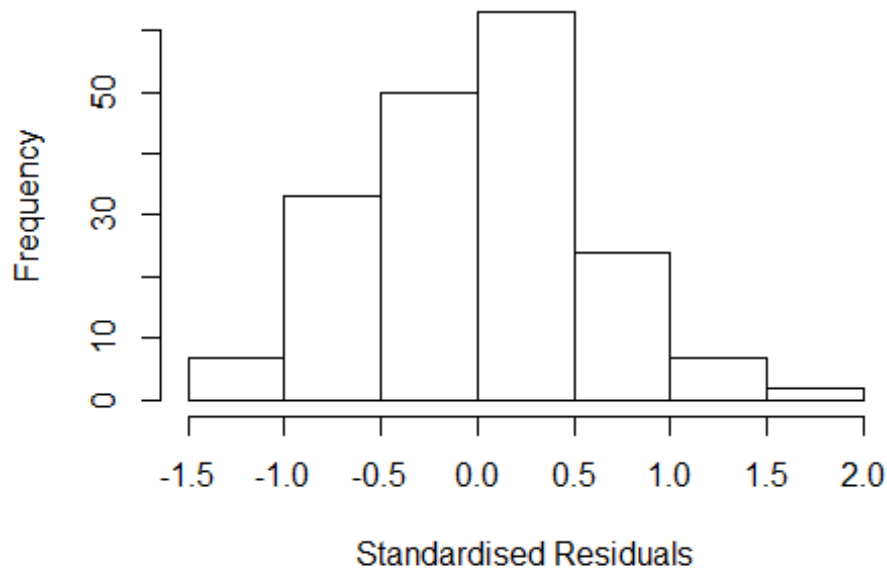
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2425 -0.4163 0.0211 0.4042 1.7906
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3494 0.0996 3.51 0.00058 ***
## FirstAuthorFemale1 -0.2489 0.1103 -2.26 0.02525 *
## LastAuthorFemale1 -0.1650 0.1054 -1.56 0.11948
## Year1997 0.5521 0.1247 4.43 1.7e-05 ***
## Year1998 0.5963 0.2698 2.21 0.02844 *
## Year1999 0.7129 0.1912 3.73 0.00026 ***
## Year2000 0.4234 0.1753 2.42 0.01677 *
## Year2001 0.3568 0.3460 1.03 0.30400
## Year2002 0.8539 0.5090 1.68 0.09531 .
## Year2003 0.7981 0.2630 3.03 0.00279 **
## Year2004 0.9460 0.2071 4.57 9.5e-06 ***
## Year2005 0.7150 0.5497 1.30 0.19514
```

```

## Year2006          0.6927      0.2570      2.70  0.00774 **
## Year2007          0.8150      0.1571      5.19  6.1e-07 ***
## Year2008          1.0525      0.2340      4.50  1.3e-05 ***
## Year2009          0.8932      0.2010      4.44  1.6e-05 ***
## Year2010          0.9161      0.2369      3.87  0.00016 ***
## Year2011          0.8081      0.1577      5.13  8.1e-07 ***
## Year2012          0.9248      0.1559      5.93  1.7e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.617
## Multiple R-squared:  0.16,   Adjusted R-squared:  0.0698
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 167 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.380  0.884  0.953  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      5.38e-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"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.448 1          1.203
## Year              1.448 16          1.012

```

Residuals from first author



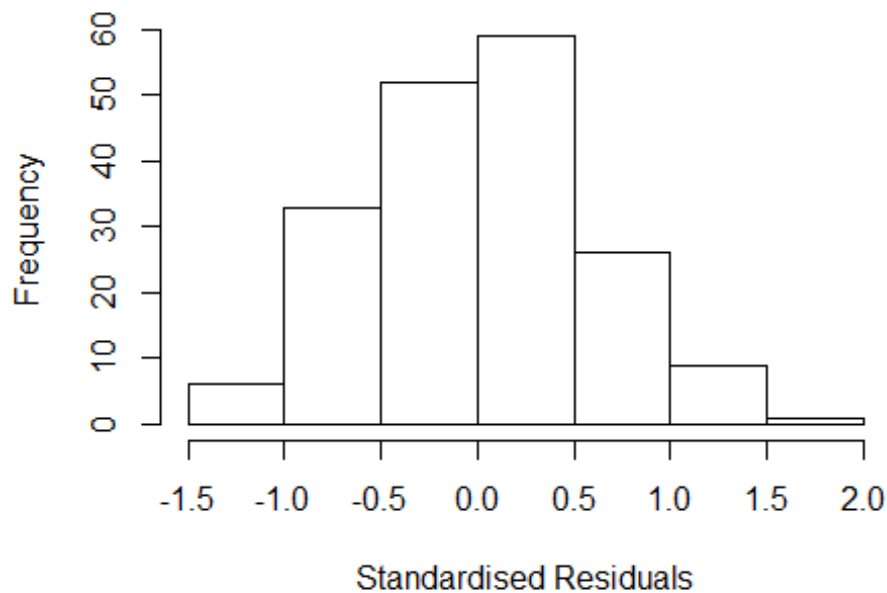
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3689 -0.3935 0.0131 0.3688 1.6656
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.321 0.102 3.15 0.00194 **
## FirstAuthorFemale1 -0.307 0.101 -3.03 0.00279 **
## Year1997 0.489 0.140 3.49 0.00062 ***
## Year1998 0.569 0.266 2.14 0.03394 *
## Year1999 0.680 0.205 3.31 0.00113 **
## Year2000 0.402 0.184 2.19 0.03002 *
## Year2001 0.355 0.329 1.08 0.28154
## Year2002 0.883 0.601 1.47 0.14392
## Year2003 0.796 0.266 2.99 0.00318 **
## Year2004 0.924 0.213 4.33 2.6e-05 ***
## Year2005 0.703 0.519 1.36 0.17688
## Year2006 0.641 0.255 2.52 0.01283 *
```

```

## Year2007          0.793      0.158      5.01  1.4e-06 ***
## Year2008          1.048      0.239      4.38  2.1e-05 ***
## Year2009          0.869      0.209      4.16  5.1e-05 ***
## Year2010          0.857      0.219      3.92  0.00013 ***
## Year2011          0.792      0.162      4.90  2.3e-06 ***
## Year2012          0.904      0.158      5.71  5.0e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.149, Adjusted R-squared:  0.0623
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.456  0.889   0.957   0.915   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      5.38e-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"
##      GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.255 1      1.120
## Year              1.255 16      1.007

```

Residuals from last author



```
## [1] "List of 0 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
## -1.1636 -0.4156 0.0186 0.4007 1.9524
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.3401 0.0986 3.45 0.00071 ***
## LastAuthorFemale1 -0.2464 0.0974 -2.53 0.01237 *
## Year1997 0.5341 0.1596 3.35 0.00101 **
## Year1998 0.5182 0.2609 1.99 0.04862 *
## Year1999 0.6311 0.1922 3.28 0.00125 **
## Year2000 0.3587 0.1576 2.28 0.02408 *
## Year2001 0.3195 0.3184 1.00 0.31703
## Year2002 0.7851 0.4034 1.95 0.05330 .
## Year2003 0.6770 0.2534 2.67 0.00828 **
## Year2004 0.9391 0.2153 4.36 2.2e-05 ***
## Year2005 0.6438 0.6993 0.92 0.35851
## Year2006 0.5708 0.2747 2.08 0.03923 *
```

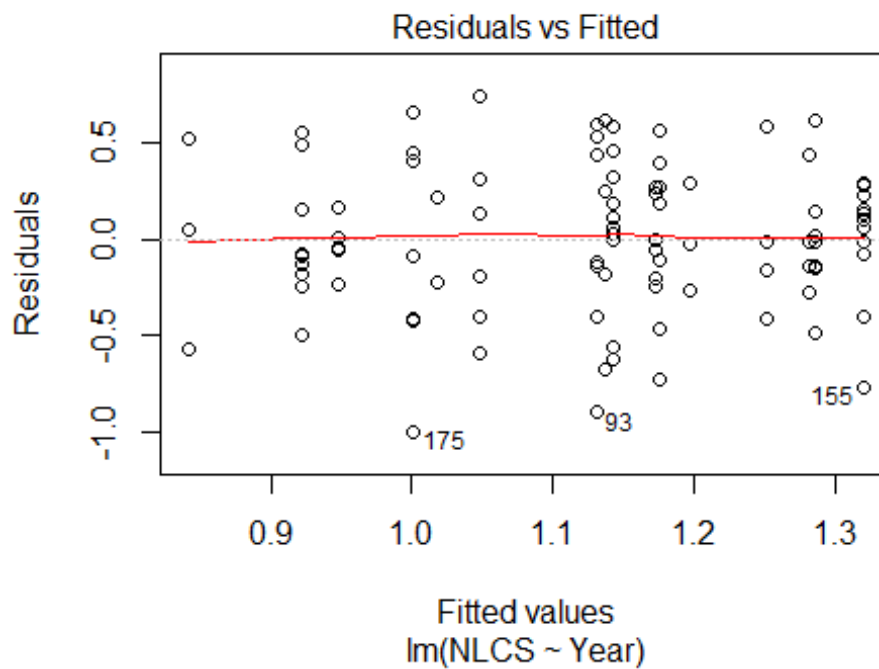


```

## Year2007          0.7467      0.1620      4.61  7.9e-06 ***
## Year2008          0.9024      0.2017      4.47  1.4e-05 ***
## Year2009          0.8235      0.1923      4.28  3.1e-05 ***
## Year2010          0.8147      0.2456      3.32  0.00112 **
## Year2011          0.7310      0.1512      4.83  3.0e-06 ***
## Year2012          0.8092      0.1545      5.24  4.8e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.135, Adjusted R-squared:  0.048
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 170 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.297  0.885   0.953   0.913   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      5.38e-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: 186"
## [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
##   12  14  16   6  13  13  16   9  11   5  19  15  12  13  11
## 2011 2012
##   15  15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2   6   4   4   0   3   7   4   6   3  11   8   6  11   9
## 2011 2012

```

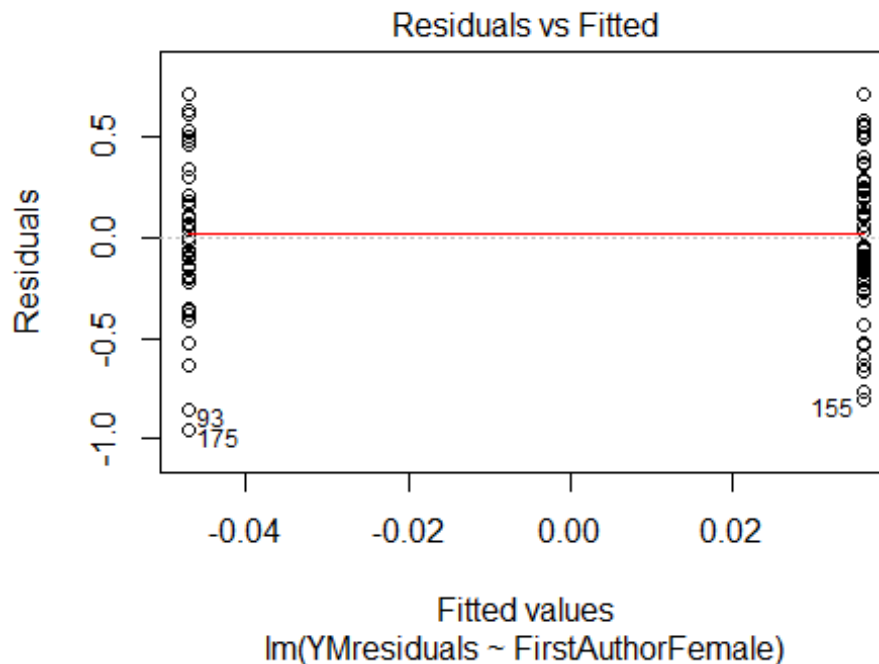
```
##      8      7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      5      2      4      0      3      6      3      5      3      9      6      6      10      8
## 2011 2012
##      8      6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 15, df = 15, p-value = 0.5
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.054, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 3.72791927319135"
## [1] "Male first author team size 2018 geometric mean: 7"
## 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, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.43754385517496"
## [1] "Male last author team size 2018 geometric mean: 4.43016905559275"

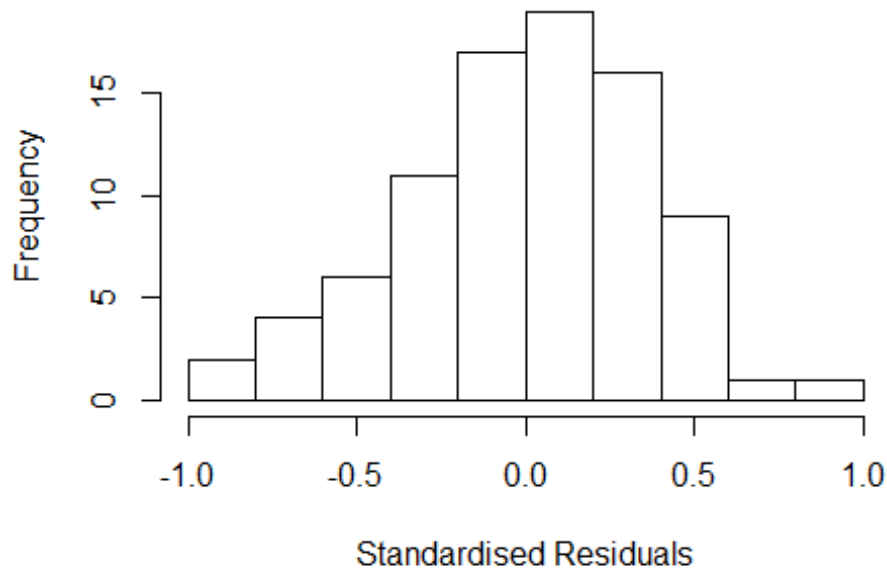
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 8.5, 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	2.940	1	1.715
LastAuthorFemale	3.581	1	1.892
UniqueAuthors	165.056	4	1.893
Year	369.769	15	1.218

Residuals from first and last author and team size



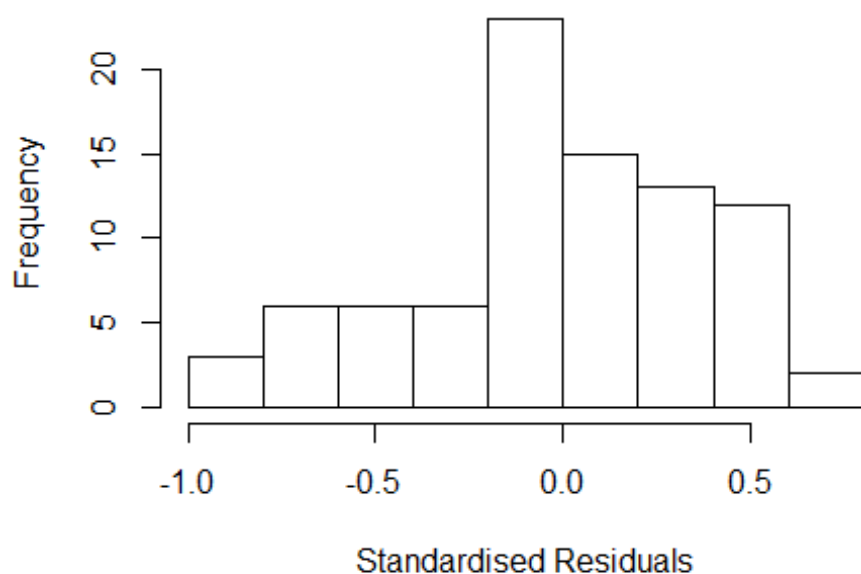
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8834 -0.2387 0.0231 0.2476 0.9414
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7303 0.1738 4.20 8.4e-05 ***
## FirstAuthorFemale1 0.0599 0.1009 0.59 0.555
## LastAuthorFemale1 0.0780 0.1308 0.60 0.553
## UniqueAuthors2 0.3765 0.2140 1.76 0.083 .
## UniqueAuthors3 0.5532 0.2162 2.56 0.013 *
## UniqueAuthors4 0.6250 0.2601 2.40 0.019 *
## UniqueAuthors5 0.5729 0.2173 2.64 0.011 *
## Year1997 -0.0782 0.2332 -0.34 0.739
## Year1998 -0.3860 0.2023 -1.91 0.061 .
## Year1999 -0.0477 0.3596 -0.13 0.895
```

```

## Year2001          -0.1491      0.2431   -0.61    0.542
## Year2002          -0.3161      0.3005   -1.05    0.297
## Year2003           0.0220      0.2289    0.10    0.924
## Year2004          -0.2203      0.1655   -1.33    0.188
## Year2005          -0.2892      0.1955   -1.48    0.144
## Year2006           0.0129      0.1779    0.07    0.942
## Year2007          -0.2772      0.2421   -1.14    0.257
## Year2008          -0.3151      0.3407   -0.92    0.359
## Year2009          -0.1478      0.1972   -0.75    0.457
## Year2010          -0.4179      0.1963   -2.13    0.037 *
## Year2011          -0.1320      0.1862   -0.71    0.481
## Year2012          -0.0295      0.2151   -0.14    0.891
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.377
## Multiple R-squared:  0.225, Adjusted R-squared:  -0.029
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 80 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.512  0.873   0.959   0.916   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      1.16e-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 2.766 1 1.663
## LastAuthorFemale 2.679 1 1.637
## Year 5.967 15 1.061

```

Residuals from first and last author



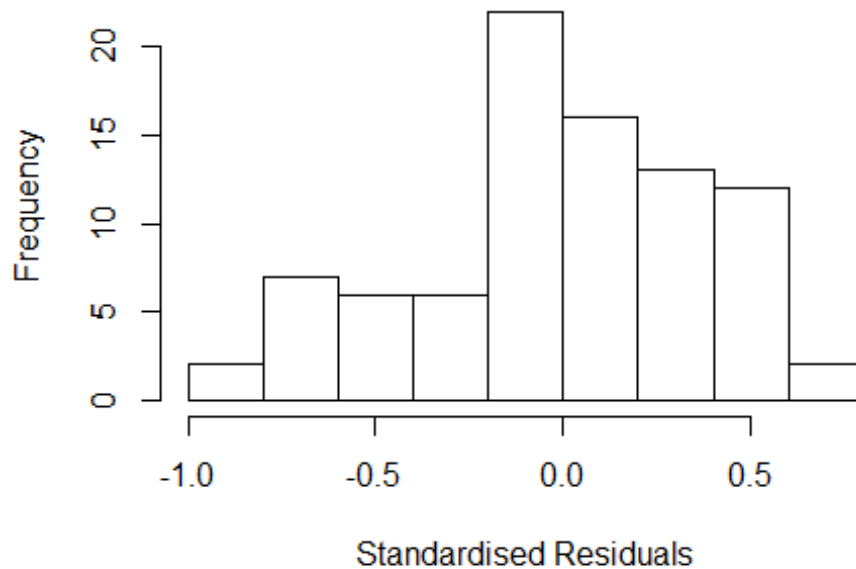
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9784 -0.1921 -0.0165 0.2375 0.7158
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.89672 0.21791 4.12 0.00011 ***
## FirstAuthorFemale1 0.11295 0.09986 1.13 0.26199
## LastAuthorFemale1 0.01566 0.11616 0.13 0.89314
## Year1997 0.09855 0.22218 0.44 0.65877
## Year1998 0.00547 0.18587 0.03 0.97659
## Year1999 0.22853 0.33927 0.67 0.50285
## Year2001 0.25789 0.24351 1.06 0.29333
## Year2002 0.10273 0.28513 0.36 0.71976
## Year2003 0.30788 0.26999 1.14 0.25815
## Year2004 0.12981 0.19018 0.68 0.49721
## Year2005 -0.07837 0.34410 -0.23 0.82052
## Year2006 0.36055 0.21814 1.65 0.10298
```

```

## Year2007          0.08167      0.31269      0.26  0.79473
## Year2008          0.06758      0.29478      0.23  0.81937
## Year2009          0.17484      0.24390      0.72  0.47593
## Year2010         -0.05452      0.22332     -0.24  0.80787
## Year2011          0.21884      0.23516      0.93  0.35535
## Year2012          0.31167      0.22688      1.37  0.17405
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.121, Adjusted R-squared:  -0.0994
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.542  0.863  0.969  0.921  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 2.622 1      1.619
## Year              2.622 15      1.033

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.99142 -0.19030 -0.00591 0.23534 0.71147
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.90536 0.19354 4.68 1.4e-05 ***
## FirstAuthorFemale1 0.11214 0.09816 1.14 0.257
## Year1997 0.08991 0.19834 0.45 0.652
## Year1998 0.00507 0.18135 0.03 0.978
## Year1999 0.22414 0.33465 0.67 0.505
## Year2001 0.24942 0.22604 1.10 0.274
## Year2002 0.10064 0.28295 0.36 0.723
## Year2003 0.29957 0.25651 1.17 0.247
## Year2004 0.12823 0.18358 0.70 0.487
## Year2005 -0.08626 0.33263 -0.26 0.796
## Year2006 0.35545 0.20058 1.77 0.081 .
## Year2007 0.08606 0.30624 0.28 0.780
```



```

## Year2008          0.06403    0.28799    0.22    0.825
## Year2009          0.17533    0.24005    0.73    0.468
## Year2010         -0.05457    0.21921   -0.25    0.804
## Year2011          0.21892    0.23087    0.95    0.346
## Year2012          0.30670    0.22018    1.39    0.168
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.401
## Multiple R-squared:  0.12,   Adjusted R-squared:  -0.0837
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 81 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.521  0.855  0.968  0.919  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 2.214  1          1.488
## Year              2.214 15          1.027

## [1] "List of 0 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
## -1.0270 -0.2030 -0.0155  0.2700  0.7642
##

```

```

## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0122    0.1794   5.64 3.4e-07 ***
## LastAuthorFemale1  0.0106    0.1145   0.09  0.93
## Year1997         -0.0169    0.1846  -0.09  0.93
## Year1998         -0.0510    0.1917  -0.27  0.79
## Year1999          0.1348    0.3417   0.39  0.69
## Year2001          0.1836    0.2265   0.81  0.42
## Year2002          0.0289    0.2848   0.10  0.92
## Year2003          0.2631    0.2654   0.99  0.32
## Year2004          0.1069    0.1898   0.56  0.58
## Year2005         -0.1611    0.3574  -0.45  0.65
## Year2006          0.3146    0.2105   1.49  0.14
## Year2007          0.0148    0.3141   0.05  0.96
## Year2008          0.0166    0.2746   0.06  0.95
## Year2009          0.1324    0.2263   0.59  0.56
## Year2010         -0.0801    0.2116  -0.38  0.71
## Year2011          0.1757    0.2342   0.75  0.46
## Year2012          0.2884    0.2197   1.31  0.19
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.427
## Multiple R-squared:  0.103, Adjusted R-squared:  -0.105
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 77 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.542  0.877  0.962  0.923  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      1.16e-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: 86"
## [1] ""
## [1] ""
## [1] "#####"

```

```

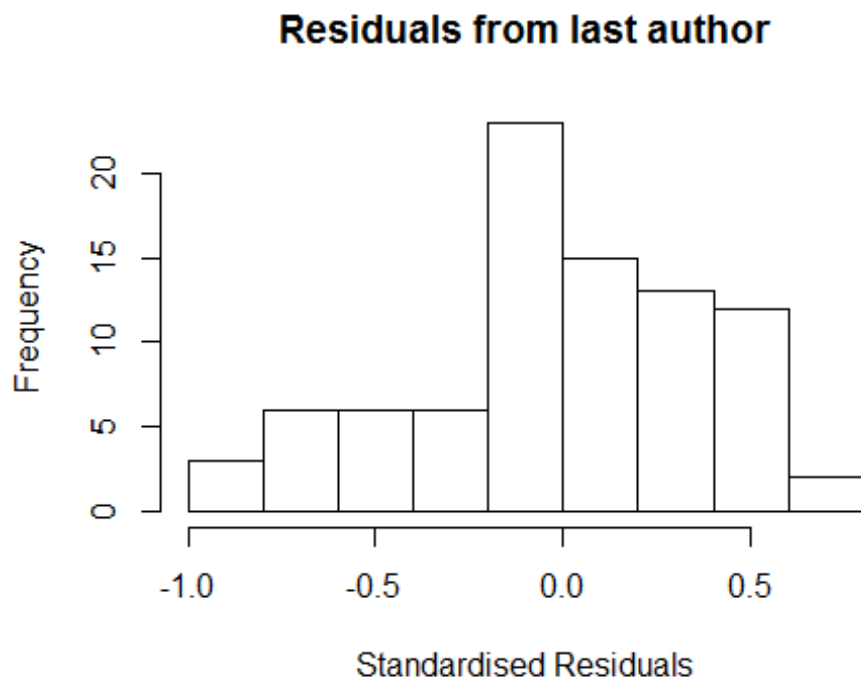
## [1] "Analysis of AJSC 2744"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996
##    3
##
## 1996
##    3
##
## 1996
##    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 2"
## [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
##    6    6    3    9    2    3    1    4    7   10    5   11    7   20   12
## 2011 2012
##   22   12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    1    6    0    0    0    3    3    5    2    5    5   16    7
## 2011 2012
##   18    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    0    0    5    0    0    0    2    2    3    2    3    4   13    6
## 2011 2012
##   13    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.87298334620742"
## [1] "Male first author team size 2018 geometric mean: 5.12992784003009"

```

```
## 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 = 2, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 9"
## [1] "Male last author team size 2018 geometric mean: 3.87298334620742"

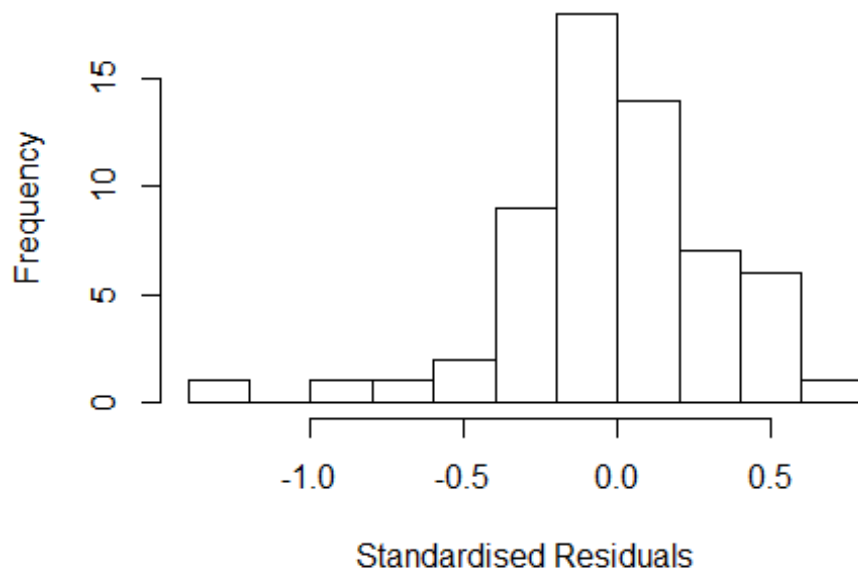
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## 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"
##
## GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 9.708 1 3.116
```

## LastAuthorFemale	5.979	1	2.445
## UniqueAuthors	58.500	4	1.663
## Year	418.404	11	1.316

Residuals from first and last author and team size



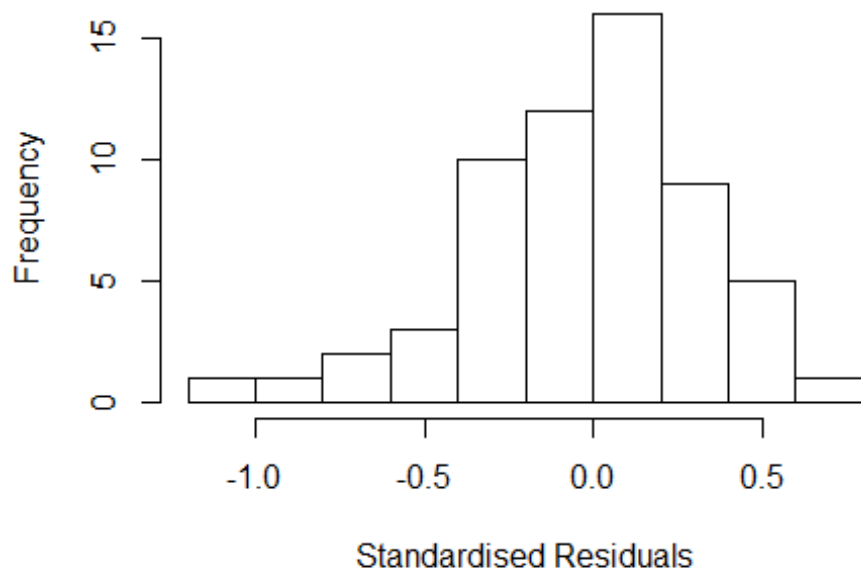
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.23803 -0.17923 -0.00919 0.19633 0.62996
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.52181 0.26361 1.98 0.0543 .
## FirstAuthorFemale1 0.43269 0.13418 3.22 0.0024 **
## LastAuthorFemale1 0.36278 0.20416 1.78 0.0828 .
## UniqueAuthors2 0.50911 0.22852 2.23 0.0313 *
## UniqueAuthors3 -0.00144 0.18866 -0.01 0.9939
## UniqueAuthors4 0.15421 0.28652 0.54 0.5933
```

```

## UniqueAuthors5      0.29819      0.26361      1.13      0.2644
## Year1999             0.40757      0.32678      1.25      0.2192
## Year2003             0.61965      0.07901      7.84      9.3e-10 ***
## Year2004             0.04913      0.35840      0.14      0.8916
## Year2005             0.03107      0.15598      0.20      0.8431
## Year2006            -0.39644      0.32139     -1.23      0.2242
## Year2007             0.10866      0.15939      0.68      0.4992
## Year2008            -0.03463      0.20924     -0.17      0.8693
## Year2009             0.15488      0.14733      1.05      0.2991
## Year2010             0.07346      0.16055      0.46      0.6496
## Year2011             0.24933      0.21886      1.14      0.2611
## Year2012            -0.07638      0.29940     -0.26      0.7999
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.321
## Multiple R-squared:  0.531, Adjusted R-squared:  0.341
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 53 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.105  0.893  0.961   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      1.67e-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 166.611 1      12.908
## LastAuthorFemale   2.676 1      1.636
## Year               347.359 11      1.305

```

Residuals from first and last author



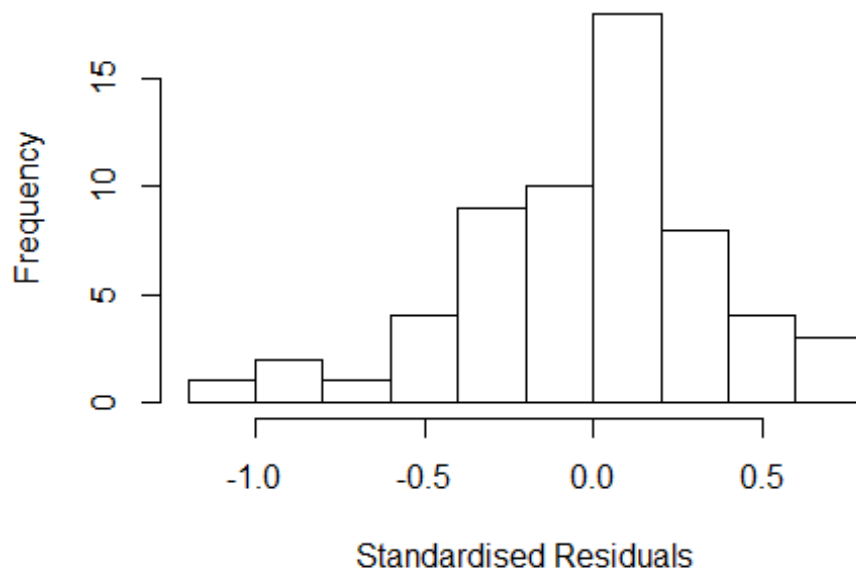
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.09312 -0.25976 0.00899 0.21169 0.62588
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.20e-01 2.91e-08 2.82e+07 < 2e-16 ***
## FirstAuthorFemale1 5.62e-01 1.10e-01 5.10e+00 6.3e-06 ***
## LastAuthorFemale1 4.91e-01 1.39e-01 3.53e+00 0.00095 ***
## Year1999 2.73e-01 1.16e-01 2.35e+00 0.02287 *
## Year2003 5.55e-01 5.53e-02 1.00e+01 3.6e-13 ***
## Year2004 5.50e-03 8.41e-01 1.00e-02 0.99481
## Year2005 -1.52e-01 1.53e-01 -9.90e-01 0.32583
## Year2006 -7.47e-01 2.90e-01 -2.58e+00 0.01319 *
## Year2007 9.28e-02 1.11e-01 8.30e-01 0.40846
## Year2008 -1.06e-01 2.04e-01 -5.20e-01 0.60535
## Year2009 -6.64e-02 1.46e-01 -4.50e-01 0.65148
## Year2010 -6.55e-02 1.39e-01 -4.70e-01 0.64036
```

```

## Year2011          -2.51e-02   9.56e-02 -2.60e-01   0.79427
## Year2012          -2.80e-01   2.45e-01 -1.14e+00   0.25854
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.379
## Multiple R-squared:  0.405, Adjusted R-squared:  0.237
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.385  0.904  0.947   0.919   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      1.67e-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 3.901  1          1.975
## Year              3.901 11          1.064

```


Residuals from first author



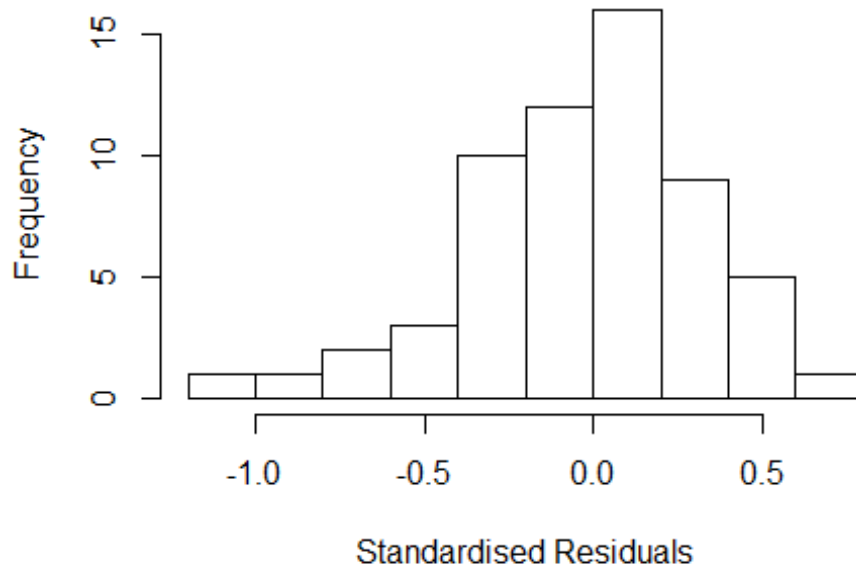
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0714 -0.2328 0.0432 0.2019 0.7798
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.20e-01 1.82e-08 4.50e+07 < 2e-16 ***
## FirstAuthorFemale1 4.14e-01 1.11e-01 3.73e+00 0.00052 ***
## Year1999 2.73e-01 1.16e-01 2.36e+00 0.02246 *
## Year2003 6.29e-01 7.36e-02 8.54e+00 4e-11 ***
## Year2004 5.50e-03 7.76e-01 1.00e-02 0.99438
## Year2005 -4.81e-02 1.98e-01 -2.40e-01 0.80862
## Year2006 -4.27e-01 1.26e-01 -3.39e+00 0.00143 **
## Year2007 3.06e-01 1.15e-01 2.67e+00 0.01050 *
## Year2008 -3.72e-02 2.10e-01 -1.80e-01 0.86061
## Year2009 6.82e-02 1.73e-01 3.90e-01 0.69566
## Year2010 8.06e-02 1.31e-01 6.20e-01 0.54002
## Year2011 1.01e-01 1.27e-01 8.00e-01 0.42929
```

```

## Year2012          4.02e-02   1.58e-01   2.50e-01   0.79993
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.303, Adjusted R-squared:  0.124
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 56 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.439  0.890  0.968  0.914  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          1.67e-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 2.61  1         1.616
## Year              2.61 11         1.045

```

Residuals from last author



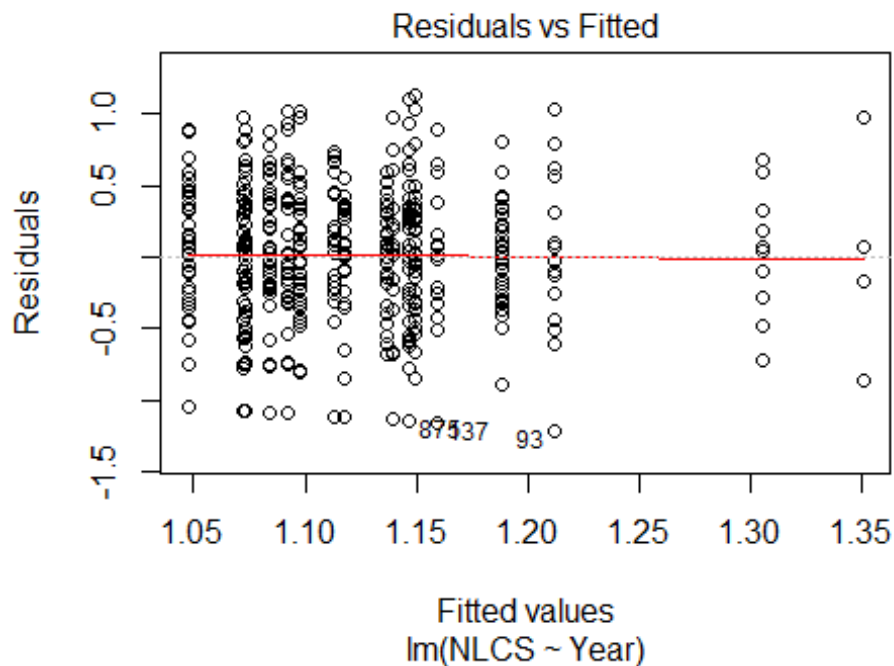
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0634 -0.3135 0.0275 0.3148 0.9285
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 8.20e-01 1.93e-08 4.25e+07 < 2e-16 ***
## LastAuthorFemale1 2.01e-01 1.29e-01 1.56e+00 0.12443
## Year1999 2.73e-01 1.15e-01 2.39e+00 0.02114 *
## Year2003 8.36e-01 2.07e-01 4.05e+00 0.00019 ***
## Year2004 5.50e-03 5.90e-01 1.00e-02 0.99260
## Year2005 2.53e-01 3.13e-01 8.10e-01 0.42263
## Year2006 -3.21e-01 4.02e-01 -8.00e-01 0.42788
## Year2007 3.77e-01 7.57e-02 4.99e+00 8.8e-06 ***
## Year2008 1.62e-01 2.55e-01 6.40e-01 0.52689
## Year2009 2.43e-01 1.38e-01 1.77e+00 0.08375 .
## Year2010 1.82e-01 1.81e-01 1.00e+00 0.32049
## Year2011 2.35e-01 1.28e-01 1.83e+00 0.07427 .
```

```

## Year2012          2.62e-01  2.13e-01  1.23e+00  0.22513
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.14,   Adjusted R-squared:  -0.0795
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.612  0.916  0.960  0.930  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           1.67e-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: 60"
## [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
##    26   34   42   41   51   38   40   44   49   44   55   56   52   59   71
## 2011 2012
##    52   55
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     4    22   18   16   14   15   28   29   33   27   39   42   34   36   47
## 2011 2012
##    36   41
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     3    19   18   12   14   13   22   25   30   20   32   35   28   31   41
## 2011 2012

```

```
## 27 35
## [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 = 0.27, df = 1, p-value = 0.6

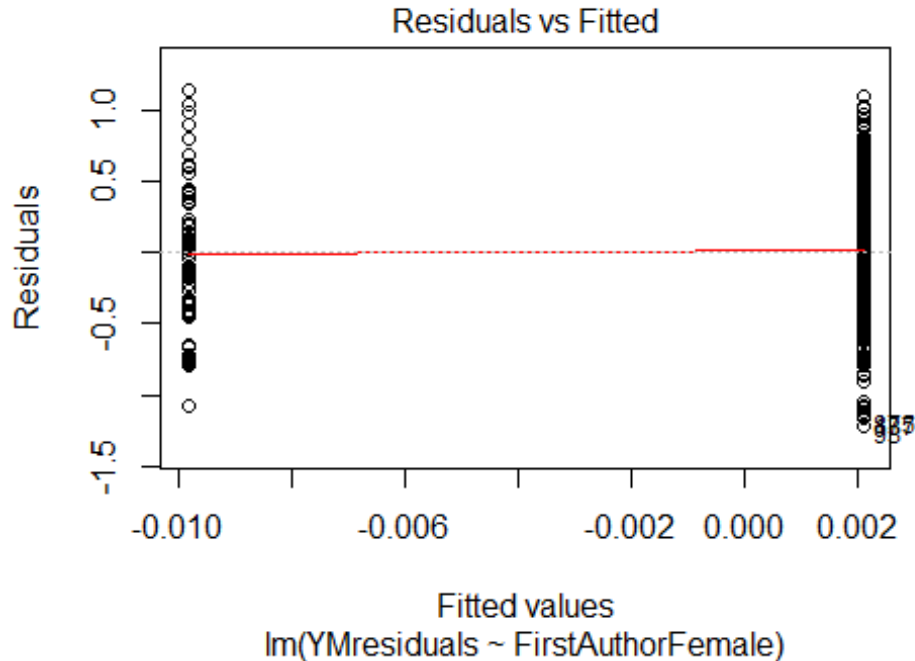
## [1] "Female first author team size 2018 geometric mean: 4.55296700002107"
## [1] "Male first author team size 2018 geometric mean: 4.18062228235863"

## 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 = 300, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
```

```
##
## [1] "Female last author team size 2018 geometric mean: 3.81457379461057"
## [1] "Male last author team size 2018 geometric mean: 4.38669363634841"

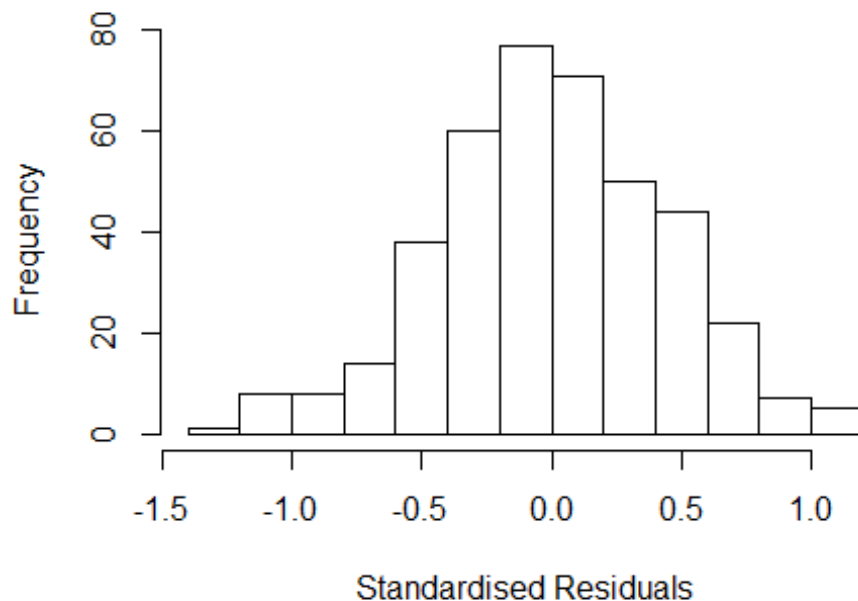
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 190, 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	1.249	1	1.117
LastAuthorFemale	1.266	1	1.125
UniqueAuthors	1.807	4	1.077
Year	2.537	16	1.030

Residuals from first and last author and team size



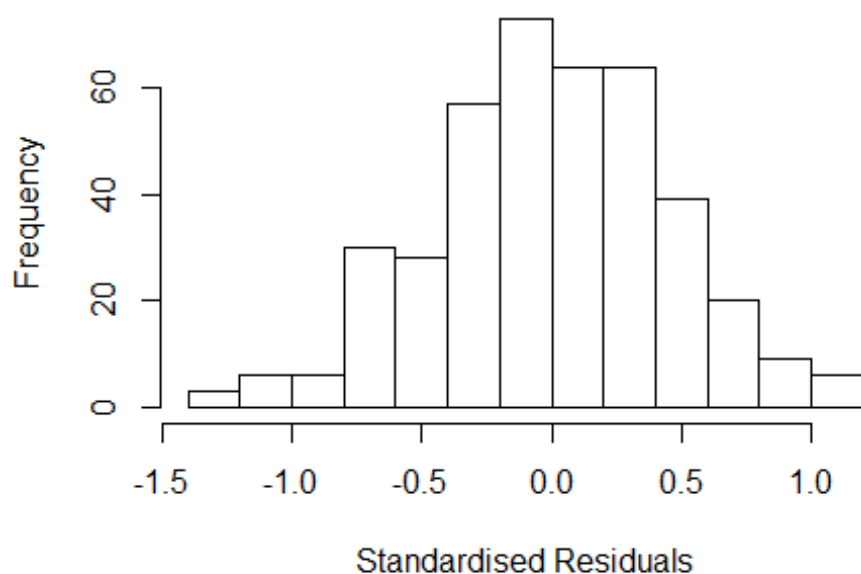
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2181 -0.2859 -0.0183  0.2912  1.1603
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2810     0.5455   2.35    0.019 *
## FirstAuthorFemale1 -0.0479     0.0581  -0.82    0.410
## LastAuthorFemale1  0.0257     0.0655   0.39    0.694
## UniqueAuthors2    -0.0946     0.1351  -0.70    0.484
## UniqueAuthors3    -0.0829     0.1374  -0.60    0.546
## UniqueAuthors4     0.0325     0.1392   0.23    0.815
## UniqueAuthors5     0.2167     0.1380   1.57    0.117
## Year1997          -0.1346     0.5442  -0.25    0.805
## Year1998          -0.0338     0.5443  -0.06    0.950
## Year1999          -0.1008     0.5392  -0.19    0.852
```

```

## Year2000      0.0543      0.5418      0.10      0.920
## Year2001     -0.2025      0.5518     -0.37      0.714
## Year2002     -0.1359      0.5368     -0.25      0.800
## Year2003     -0.0954      0.5384     -0.18      0.859
## Year2004     -0.0822      0.5323     -0.15      0.877
## Year2005     -0.1420      0.5344     -0.27      0.791
## Year2006     -0.1742      0.5344     -0.33      0.745
## Year2007     -0.2348      0.5361     -0.44      0.662
## Year2008     -0.2639      0.5325     -0.50      0.621
## Year2009     -0.2646      0.5413     -0.49      0.625
## Year2010     -0.2000      0.5366     -0.37      0.710
## Year2011     -0.3039      0.5368     -0.57      0.572
## Year2012     -0.1124      0.5402     -0.21      0.835
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.441
## Multiple R-squared:  0.0881, Adjusted R-squared:  0.0356
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 372 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.427  0.880  0.953  0.909  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      2.47e-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 1.174 1      1.084
## LastAuthorFemale  1.346 1      1.160
## Year              1.578 16      1.014

```


Residuals from first and last author



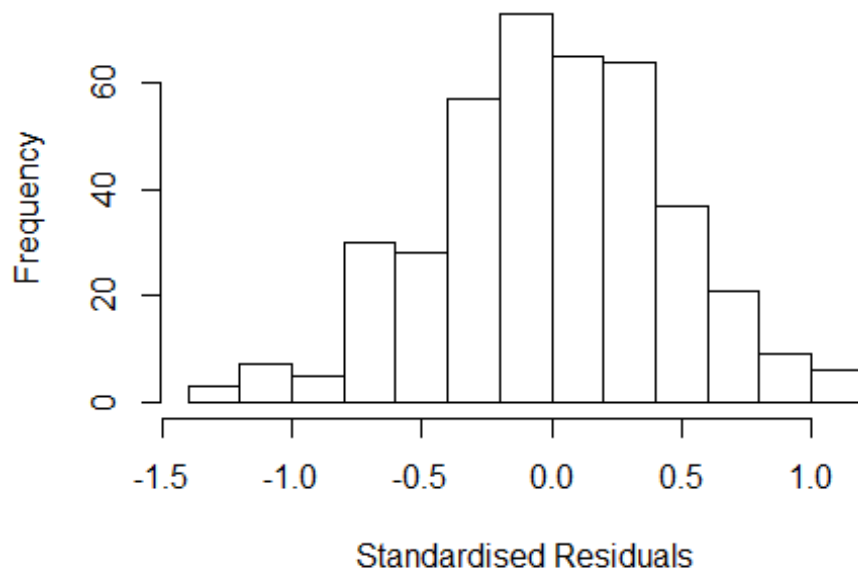
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22709 -0.29316 -0.00146 0.30774 1.11426
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.20474 0.77582 1.55 0.12
## FirstAuthorFemale1 -0.07749 0.06218 -1.25 0.21
## LastAuthorFemale1 -0.01223 0.06932 -0.18 0.86
## Year1997 -0.07627 0.78569 -0.10 0.92
## Year1998 0.02235 0.77958 0.03 0.98
## Year1999 -0.05228 0.77983 -0.07 0.95
## Year2000 0.10810 0.78052 0.14 0.89
## Year2001 -0.08685 0.78996 -0.11 0.91
## Year2002 -0.01947 0.77771 -0.03 0.98
## Year2003 -0.03712 0.78152 -0.05 0.96
## Year2004 0.00758 0.77405 0.01 0.99
## Year2005 -0.02404 0.77507 -0.03 0.98
```

```

## Year2006      -0.07681    0.77621   -0.10    0.92
## Year2007      -0.08330    0.77784   -0.11    0.91
## Year2008      -0.17255    0.77156   -0.22    0.82
## Year2009      -0.12235    0.78035   -0.16    0.88
## Year2010      -0.08494    0.77668   -0.11    0.91
## Year2011      -0.13857    0.77625   -0.18    0.86
## Year2012       0.01038    0.78021    0.01    0.99
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.0246, Adjusted R-squared:  -0.0209
## Convergence in 28 IRWLS iterations
##
## Robustness weights:
## 37 weights are ~= 1. The remaining 368 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.432  0.862  0.949  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      2.47e-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 1.173 1          1.083
## Year              1.173 16          1.005

```

Residuals from first author



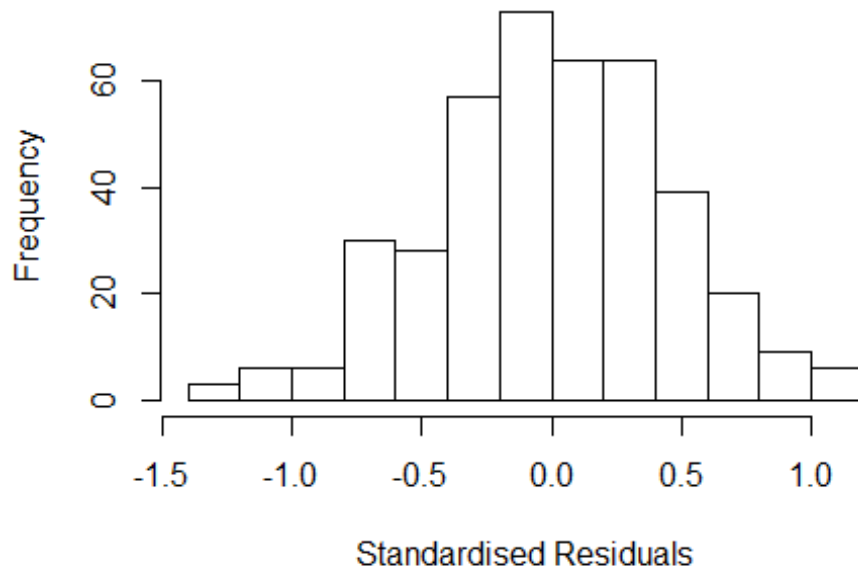
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.225138 -0.291614 -0.000398  0.308922  1.125932
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.19307    0.74543    1.60   0.11
## FirstAuthorFemale1 -0.07763    0.06221   -1.25   0.21
## Year1997         -0.06567    0.75758   -0.09   0.93
## Year1998          0.03207    0.75495    0.04   0.97
## Year1999         -0.04184    0.75301   -0.06   0.96
## Year2000          0.11733    0.75559    0.16   0.88
## Year2001         -0.07655    0.76435   -0.10   0.92
## Year2002         -0.00899    0.75042   -0.01   0.99
## Year2003         -0.02583    0.75179   -0.03   0.97
## Year2004          0.01767    0.74728    0.02   0.98
## Year2005         -0.01445    0.75011   -0.02   0.98
## Year2006         -0.06649    0.74960   -0.09   0.93
```

```

## Year2007          -0.07291    0.75038   -0.10    0.92
## Year2008          -0.16461    0.75022   -0.22    0.83
## Year2009          -0.11146    0.75147   -0.15    0.88
## Year2010          -0.07439    0.74961   -0.10    0.92
## Year2011          -0.12845    0.75006   -0.17    0.86
## Year2012           0.02034    0.75358    0.03    0.98
##
## Robust residual standard error: 0.447
## Multiple R-squared:  0.0245, Adjusted R-squared:  -0.0183
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 39 weights are ~= 1. The remaining 366 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.434  0.859  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      2.47e-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 1.33 1          1.153
## Year            1.33 16          1.009

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2174 -0.2947 -0.0109 0.3070 1.1069
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.21314 0.76263 1.59 0.11
## LastAuthorFemale1 -0.01397 0.06926 -0.20 0.84
## Year1997 -0.08690 0.77319 -0.11 0.91
## Year1998 0.00431 0.76680 0.01 1.00
## Year1999 -0.06022 0.76689 -0.08 0.94
## Year2000 0.09086 0.76683 0.12 0.91
## Year2001 -0.12920 0.77625 -0.17 0.87
## Year2002 -0.04774 0.76443 -0.06 0.95
## Year2003 -0.05349 0.76852 -0.07 0.94
## Year2004 -0.01028 0.76084 -0.01 0.99
## Year2005 -0.05822 0.76192 -0.08 0.94
## Year2006 -0.09919 0.76268 -0.13 0.90
```

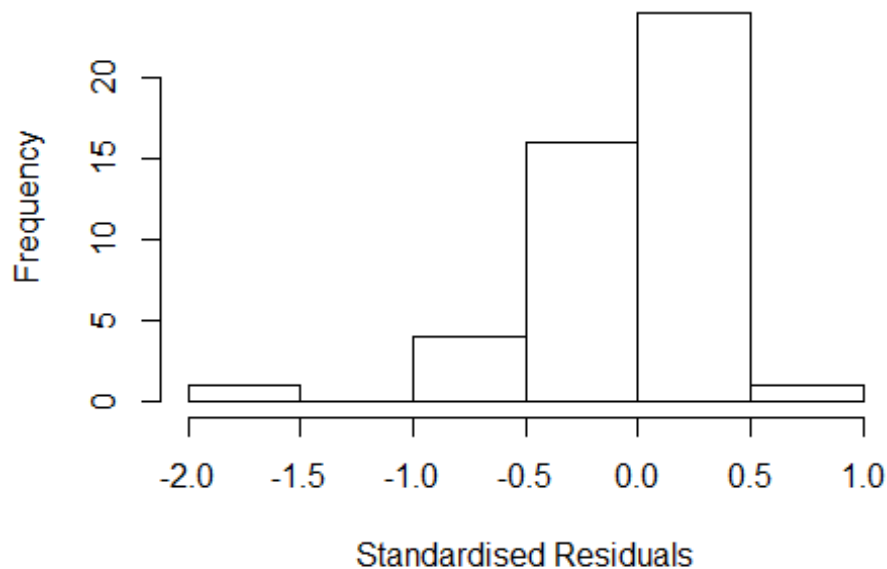
```

## Year2007          -0.10840      0.76374    -0.14      0.89
## Year2008          -0.19905      0.75867    -0.26      0.79
## Year2009          -0.14575      0.76689    -0.19      0.85
## Year2010          -0.09810      0.76336    -0.13      0.90
## Year2011          -0.16790      0.76339    -0.22      0.83
## Year2012          -0.01813      0.76630    -0.02      0.98
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.0203, Adjusted R-squared:  -0.0227
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 33 weights are ~= 1. The remaining 372 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.451  0.863  0.953  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      2.47e-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: 405"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    5    2    4    2    4    8    3    5    2    2    6    6    4    6    2
## 2012
##    7
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    1    3    0    4    6    2    3    2    2    5    4    3    5    2
## 2012
##    6
##

```

```
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    0    0    3    0    4    5    2    3    2    2    5    4    3    5    2
## 2012
##    6
## [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: 3.83365862547764"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale    2.593  1         1.610
## LastAuthorFemale    11.321  1         3.365
## UniqueAuthors      172.767  4         1.904
## Year                1815.347 12         1.367
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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"
```

```

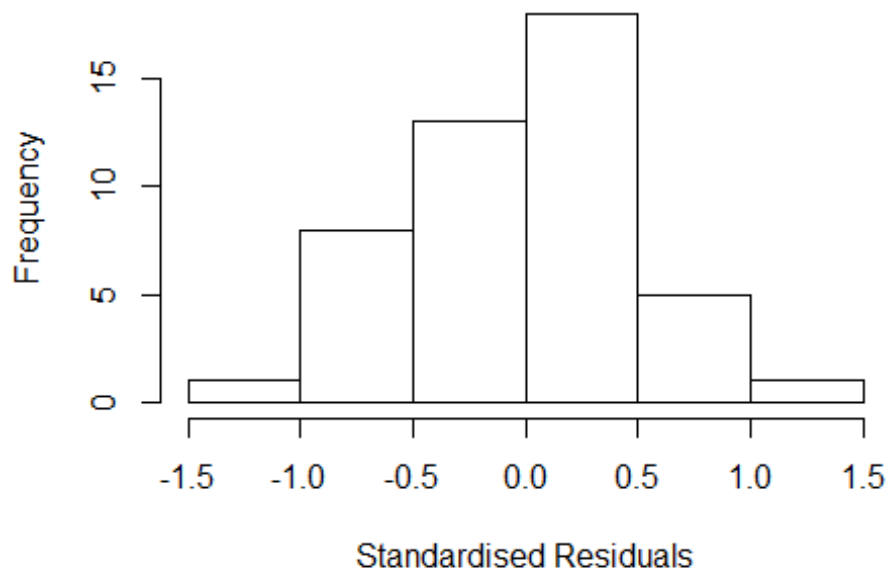
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.9535 -0.3528  0.0455  0.2471  0.8967
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.8974      0.3410   5.56 6.7e-06 ***
## FirstAuthorFemale1  0.0444      0.1630   0.27  0.7876
## LastAuthorFemale1  0.1919      0.1953   0.98  0.3345
## UniqueAuthors2     -1.1474      0.3895  -2.95  0.0066 **
## UniqueAuthors3     -1.0709      0.2922  -3.66  0.0011 **
## UniqueAuthors4     -0.7280      0.3539  -2.06  0.0494 *
## UniqueAuthors5     -0.7533      0.2163  -3.48  0.0017 **
## Year2001            0.1691      0.2600   0.65  0.5210
## Year2002           -0.0941      0.2771  -0.34  0.7367
## Year2003           -0.1397      0.2826  -0.49  0.6251
## Year2004            0.2171      0.2441   0.89  0.3818
## Year2005           -0.1039      0.3323  -0.31  0.7569
## Year2006           -0.5128      0.3428  -1.50  0.1462
## Year2007            0.1802      0.2266   0.80  0.4334
## Year2008            0.0759      0.3118   0.24  0.8094
## Year2009            0.0778      0.3884   0.20  0.8427
## Year2010            0.6391      0.3122   2.05  0.0505 .
## Year2011           -0.7892      0.3990  -1.98  0.0582 .
## Year2012           -0.4730      0.5242  -0.90  0.3748
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.499, Adjusted R-squared:  0.166
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 42 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0132 0.9050 0.9420 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          2.17e-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 2.254 1      1.501
## LastAuthorFemale 6.819 1      2.611
## Year             12.380 12      1.111
```

Residuals from first and last author



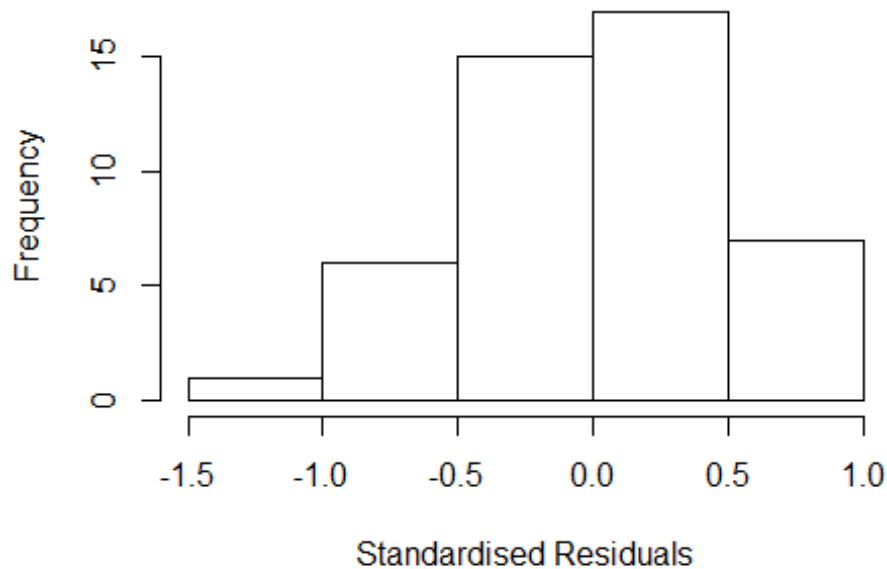
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0893 -0.3740  0.0364  0.3929  1.0179
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9003     0.2722   3.31  0.0024 **
## FirstAuthorFemale1  0.0483     0.1666   0.29  0.7740
## LastAuthorFemale1  0.3068     0.2092   1.47  0.1526
```

```

## Year2001      0.2434      0.3010      0.81      0.4249
## Year2002      0.1890      0.3685      0.51      0.6116
## Year2003      0.4807      0.4422      1.09      0.2854
## Year2004      0.2426      0.3653      0.66      0.5115
## Year2005      0.0249      0.3403      0.07      0.9422
## Year2006      0.0482      0.3134      0.15      0.8789
## Year2007      0.3022      0.3130      0.97      0.3418
## Year2008      0.3751      0.5071      0.74      0.4650
## Year2009      0.2206      0.3238      0.68      0.5008
## Year2010      0.4817      0.3655      1.32      0.1972
## Year2011     -0.1688      0.8002     -0.21      0.8343
## Year2012     -0.3252      0.4525     -0.72      0.4777
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.226, Adjusted R-squared:  -0.124
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  5 weights are ~= 1. The remaining 41 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.704  0.918  0.949  0.937  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          2.17e-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 1.995 1          1.413
## Year              1.995 12          1.029

```

Residuals from first author



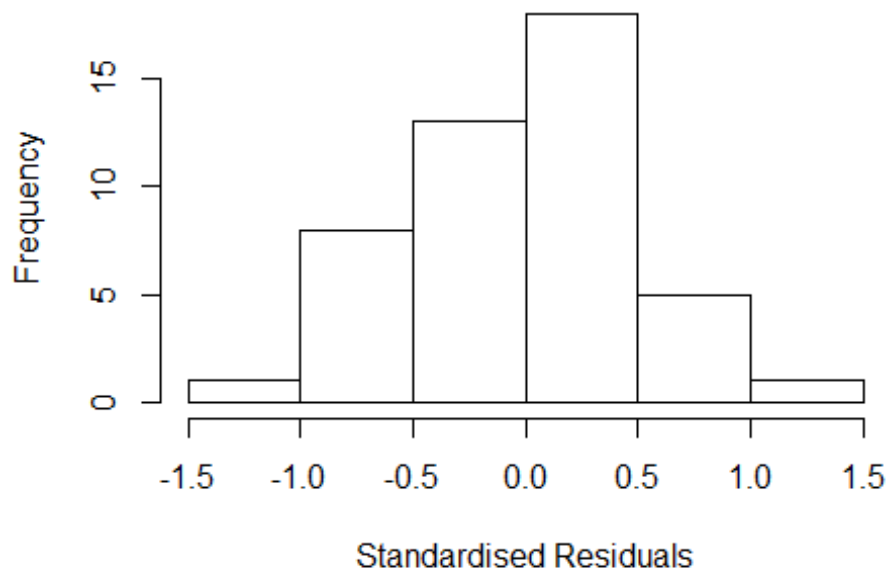
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0876 -0.3428 0.0235 0.3462 0.9017
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9966 0.1918 5.20 1.1e-05 ***
## FirstAuthorFemale1 0.0615 0.1800 0.34 0.74
## Year2001 0.1437 0.2318 0.62 0.54
## Year2002 0.0909 0.3228 0.28 0.78
## Year2003 0.3844 0.3993 0.96 0.34
## Year2004 0.1462 0.3098 0.47 0.64
## Year2005 0.2354 0.1983 1.19 0.24
## Year2006 0.0986 0.2964 0.33 0.74
## Year2007 0.3831 0.2498 1.53 0.13
## Year2008 0.3635 0.4018 0.90 0.37
## Year2009 0.1196 0.2608 0.46 0.65
## Year2010 0.4532 0.3110 1.46 0.15
```

```

## Year2011          -0.2651      0.7855   -0.34      0.74
## Year2012          -0.3053      0.4763   -0.64      0.53
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.573
## Multiple R-squared:  0.188, Adjusted R-squared:  -0.142
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.699  0.882  0.965  0.932  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      2.17e-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 6.614 1          2.572
## Year            6.614 12          1.082

```

Residuals from last author



```
## [1] "List of 0 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
## -1.1065 -0.3879  0.0489  0.3803  1.0034
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8994     0.2749   3.27  0.0026 **
## LastAuthorFemale1 0.3108     0.2118   1.47  0.1521
## Year2001        0.2571     0.2990   0.86  0.3962
## Year2002        0.2071     0.3597   0.58  0.5688
## Year2003        0.4816     0.4466   1.08  0.2889
## Year2004        0.2432     0.3682   0.66  0.5136
## Year2005        0.0218     0.3449   0.06  0.9499
## Year2006        0.0712     0.3119   0.23  0.8208
## Year2007        0.3106     0.3127   0.99  0.3281
## Year2008        0.4004     0.4943   0.81  0.4238
## Year2009        0.2387     0.3181   0.75  0.4586
## Year2010        0.5034     0.3604   1.40  0.1721
```

```

## Year2011          -0.1679      0.8264   -0.20   0.8403
## Year2012          -0.3098      0.4489   -0.69   0.4951
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.561
## Multiple R-squared:  0.226, Adjusted R-squared:  -0.0884
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.677  0.912  0.947   0.934   0.992   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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: 46"
## [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 2008 2009 2010 2011
##    6   10    8    5    6    5    5    7    2    4    7    6    4    4    2
## 2012
##    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2008 2009 2010 2011
##    2    4    2    2    2    1    3    3    1    3    4    4    4    1    2
## 2012
##    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2008 2009 2010 2011
##    2    4    2    1    1    1    3    3    1    2    3    4    4    1    2

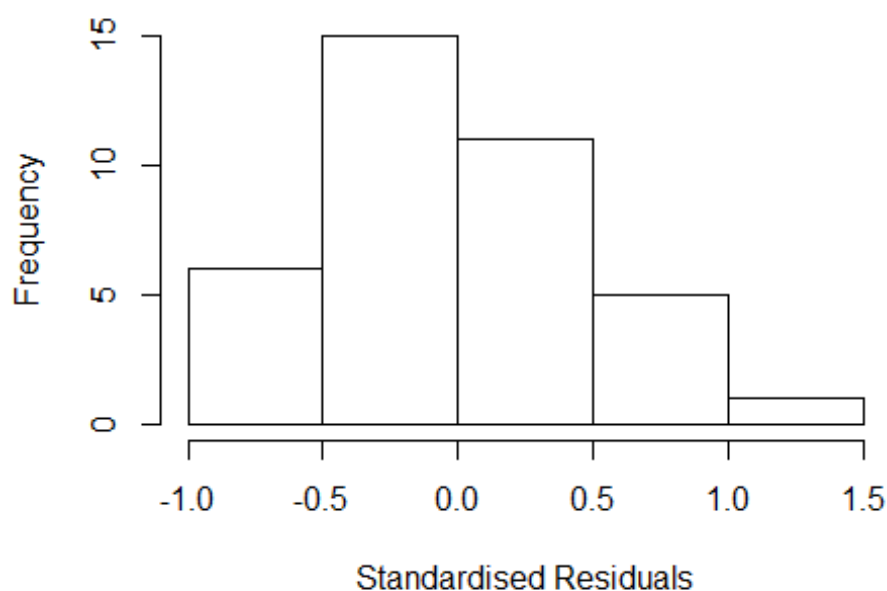
```

```

## 2012
## 4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 6"
## [1] "Male first author team size 2018 geometric mean: 4.71769398031653"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 6"
## [1] "Male last author team size 2018 geometric mean: 4.71769398031653"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2, 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"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 8.256e+14 1      2.873e+07
## LastAuthorFemale  8.230e+14 1      2.869e+07
## Year              5.121e+15 15      3.339e+00

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -8.85e-01 -3.25e-01 1.67e-16 2.22e-01 1.32e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.4370 0.0803 30.37 < 2e-16 ***
## FirstAuthorFemale1 0.1058 0.2288 0.46 0.64886
## LastAuthorFemale1 0.0494 0.2251 0.22 0.82858
## Year1997 -1.8659 0.2118 -8.81 2.5e-08 ***
## Year1998 -1.0430 0.8829 -1.18 0.25131
## Year1999 -1.2530 0.0803 -15.61 1.1e-12 ***
## Year2000 -1.1610 0.0803 -14.47 4.7e-12 ***
## Year2001 -0.5770 0.0803 -7.19 5.8e-07 ***
## Year2002 -1.3049 0.2682 -4.87 9.4e-05 ***
## Year2003 -1.4152 0.1781 -7.94 1.3e-07 ***
## Year2004 -0.5740 0.0803 -7.15 6.3e-07 ***
## Year2005 -1.7094 0.2984 -5.73 1.3e-05 ***
```

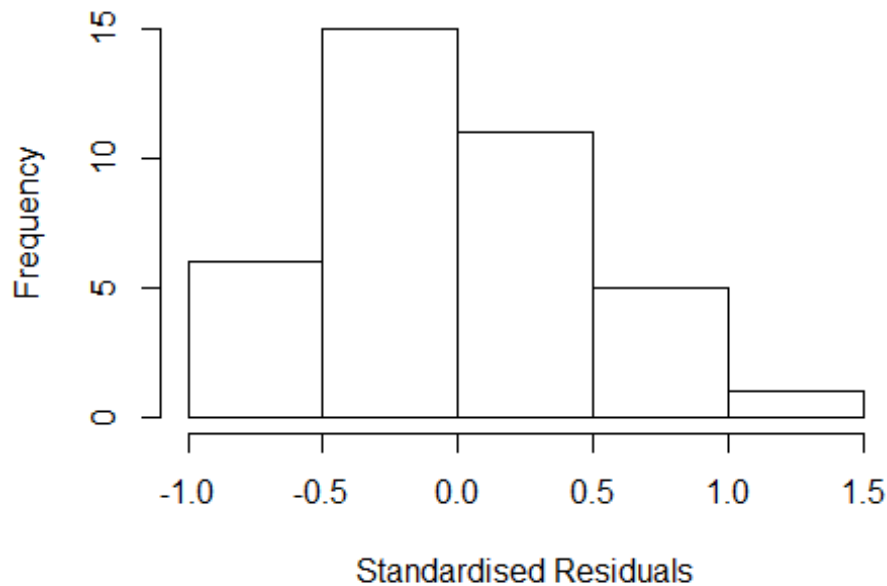


```

## Year2006          -0.9992      0.4376   -2.28  0.03347 *
## Year2008          -0.8058      0.1801   -4.48  0.00023 ***
## Year2009          -1.0874      0.2375   -4.58  0.00018 ***
## Year2010          -1.1592      0.2825   -4.10  0.00055 ***
## Year2011          -1.8795      0.4500   -4.18  0.00047 ***
## Year2012          -1.3010      0.5304   -2.45  0.02348 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.729
## Multiple R-squared:  0.466, Adjusted R-squared:  0.0115
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.724  0.943  0.963  0.951  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.63e-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 -8.749e+14 1      NaN
## Year              -8.749e+14 15      NaN

```

Residuals from first author



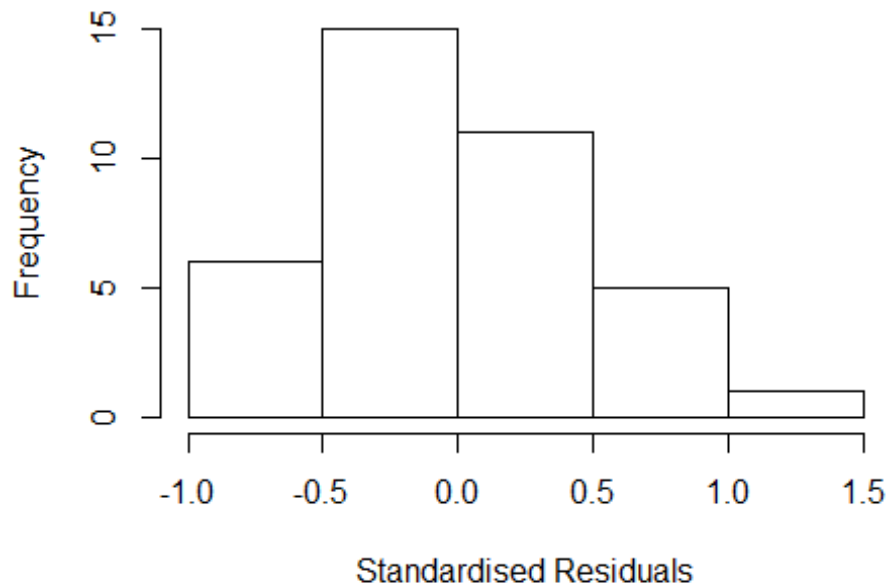
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.885 -0.320 0.000 0.245 1.331
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.4370 0.0803 30.35 < 2e-16 ***
## FirstAuthorFemale1 0.1163 0.2268 0.51 0.61345
## Year1997 -1.8656 0.2124 -8.79 1.8e-08 ***
## Year1998 -1.0430 0.9247 -1.13 0.27209
## Year1999 -1.2530 0.0803 -15.61 5.0e-13 ***
## Year2000 -1.1610 0.0803 -14.46 2.2e-12 ***
## Year2001 -0.5770 0.0803 -7.19 4.4e-07 ***
## Year2002 -1.3054 0.2695 -4.84 8.7e-05 ***
## Year2003 -1.4060 0.1920 -7.32 3.3e-07 ***
## Year2004 -0.5740 0.0803 -7.15 4.8e-07 ***
## Year2005 -1.7147 0.3030 -5.66 1.3e-05 ***
## Year2006 -0.9990 0.4438 -2.25 0.03524 *
```

```

## Year2008          -0.8084      0.1798    -4.50  0.00020 ***
## Year2009          -1.0745      0.2067    -5.20  3.8e-05 ***
## Year2010          -1.1203      0.2406    -4.66  0.00014 ***
## Year2011          -1.8795      0.4558    -4.12  0.00048 ***
## Year2012          -1.3146      0.5374    -2.45  0.02332 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.693
## Multiple R-squared:  0.469, Adjusted R-squared:  0.0643
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 28 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.693  0.938  0.959  0.946  0.993  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.63e-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 1.15e+16 1      1.072e+08
## Year              1.15e+16 15      3.430e+00

```

Residuals from last author



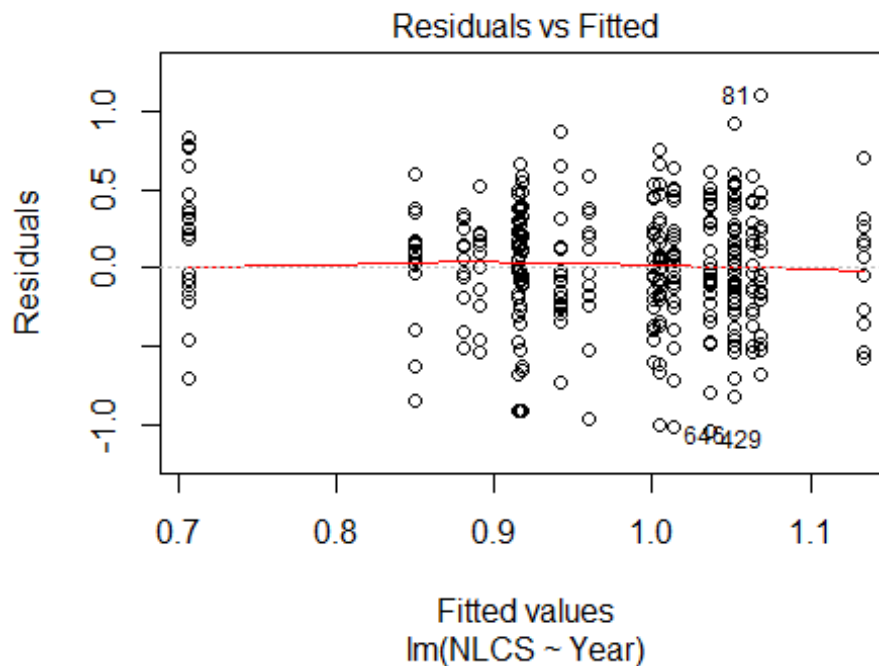
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.85e-01 -2.90e-01 -4.44e-16 2.63e-01 1.29e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.4370 0.0803 30.35 < 2e-16 ***
## LastAuthorFemale1 0.0747 0.2286 0.33 0.74714
## Year1997 -1.8657 0.2123 -8.79 1.8e-08 ***
## Year1998 -1.0430 0.9162 -1.14 0.26775
## Year1999 -1.2530 0.0803 -15.61 5.0e-13 ***
## Year2000 -1.1610 0.0803 -14.46 2.2e-12 ***
## Year2001 -0.5770 0.0803 -7.19 4.4e-07 ***
## Year2002 -1.3053 0.2693 -4.85 8.6e-05 ***
## Year2003 -1.3532 0.1346 -10.06 1.8e-09 ***
## Year2004 -0.5740 0.0803 -7.15 4.8e-07 ***
## Year2005 -1.6565 0.2363 -7.01 6.4e-07 ***
## Year2006 -0.9991 0.4426 -2.26 0.03477 *
```

```

## Year2008          -0.7783      0.1631    -4.77  0.00010 ***
## Year2009          -1.0941      0.2388    -4.58  0.00016 ***
## Year2010          -1.0787      0.2423    -4.45  0.00022 ***
## Year2011          -1.8795      0.4546    -4.13  0.00047 ***
## Year2012          -1.2712      0.4923    -2.58  0.01739 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.7
## Multiple R-squared:  0.466, Adjusted R-squared:  0.0585
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 29 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.716  0.940  0.963  0.948  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      2.63e-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: 38"
## [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
##   35   30   31   31   34   30   39   38   38   32   45   57   27   21   35
## 2011 2012
##   30   28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   15   16   13   11   17   16   27   26   24   20   29   44   20   12   24
## 2011 2012
##   19   21

```

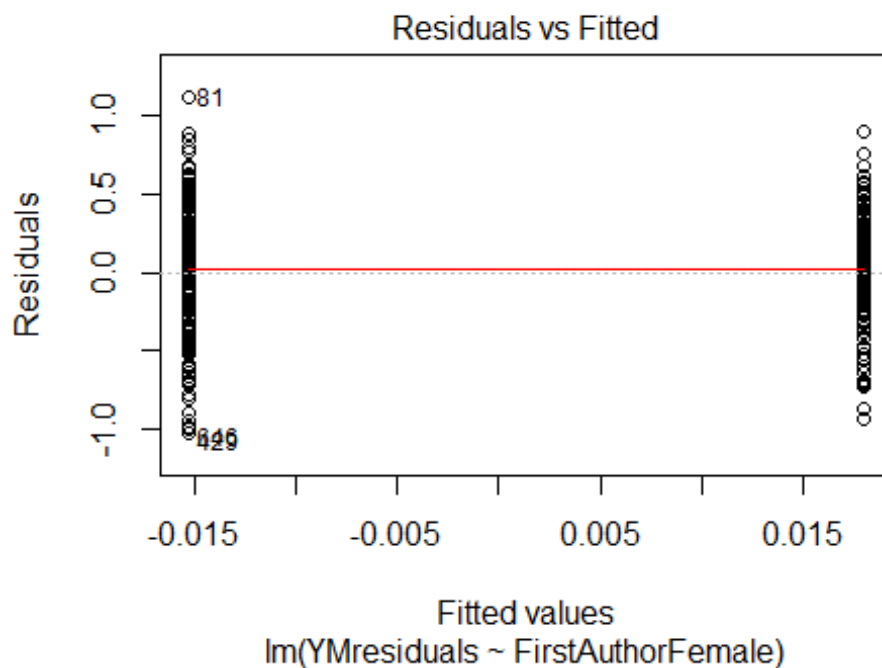
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 13 11 10 15 13 22 24 19 17 27 39 19 10 19
## 2011 2012
## 16 19
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.7
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 8.1, df = 1, p-value = 0.004
## [1] "Female first author team size 2018 geometric mean: 4.55573246476705"
## [1] "Male first author team size 2018 geometric mean: 3.79445446885362"
## 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 = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.10714661365223"
## [1] "Male last author team size 2018 geometric mean: 4.12443538891977"

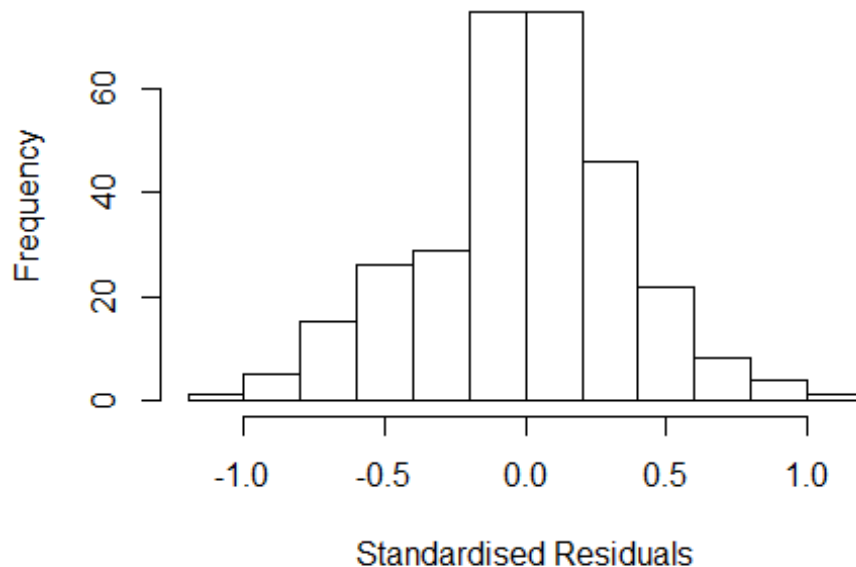
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 76, 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	1.900	1	1.378
LastAuthorFemale	1.372	1	1.171
UniqueAuthors	5.811	4	1.246
Year	8.204	16	1.068

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.08985 -0.19046 0.00715 0.22938 1.03228
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8885 0.1465 6.06 4.2e-09 ***
## FirstAuthorFemale1 0.0222 0.0526 0.42 0.67360
## LastAuthorFemale1 -0.0736 0.0519 -1.42 0.15732
## UniqueAuthors2 0.3934 0.1020 3.86 0.00014 ***
## UniqueAuthors3 0.2811 0.1008 2.79 0.00565 **
## UniqueAuthors4 0.2783 0.1085 2.57 0.01083 *
## UniqueAuthors5 0.3567 0.0980 3.64 0.00032 ***
## Year1997 -0.1462 0.1865 -0.78 0.43385
## Year1998 -0.0798 0.1642 -0.49 0.62747
## Year1999 -0.2144 0.1467 -1.46 0.14506
```

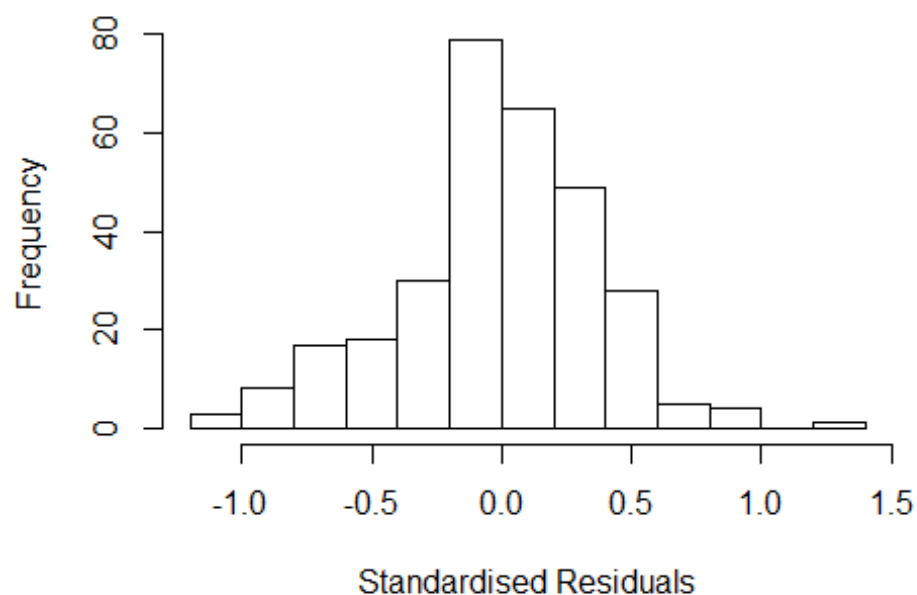


```

## Year2000          -0.3240      0.1381    -2.35  0.01965 *
## Year2001          -0.0451      0.1681    -0.27  0.78879
## Year2002          -0.2101      0.1320    -1.59  0.11254
## Year2003          -0.4599      0.1509    -3.05  0.00252 **
## Year2004          -0.3232      0.1372    -2.36  0.01916 *
## Year2005          -0.2824      0.1347    -2.10  0.03691 *
## Year2006          -0.1101      0.1307    -0.84  0.40030
## Year2007          -0.1167      0.1257    -0.93  0.35406
## Year2008          -0.3341      0.1360    -2.46  0.01465 *
## Year2009          -0.3080      0.1444    -2.13  0.03381 *
## Year2010          -0.1166      0.1482    -0.79  0.43207
## Year2011          -0.1241      0.1424    -0.87  0.38423
## Year2012          -0.2626      0.1504    -1.75  0.08190 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.198, Adjusted R-squared:  0.136
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 278 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.288  0.862  0.951   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      3.26e-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 1.338 1          1.157
## LastAuthorFemale  1.309 1          1.144
## Year              1.747 16          1.018

```

Residuals from first and last author



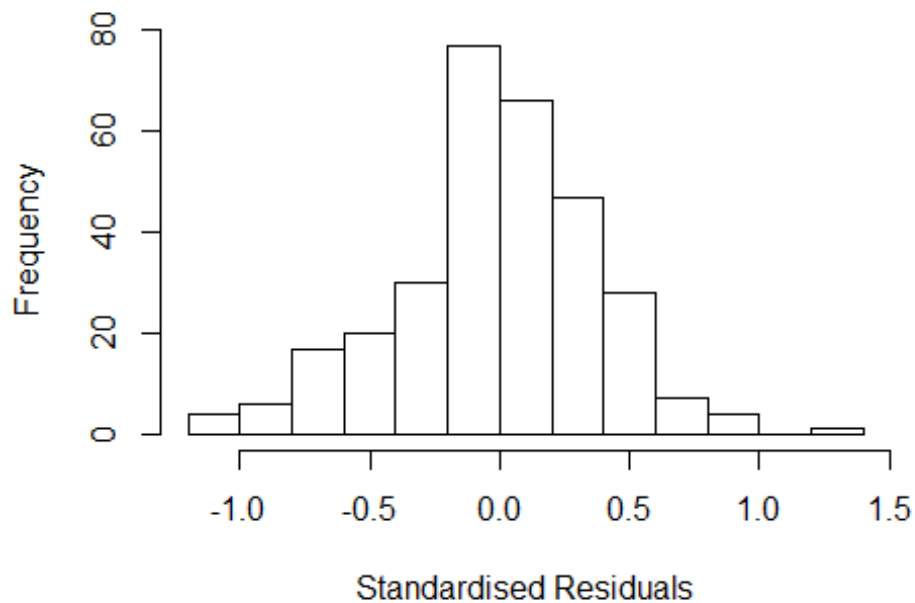
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.04975 -0.19728 -0.00242 0.23273 1.24859
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1526 0.1004 11.48 <2e-16 ***
## FirstAuthorFemale1 0.0779 0.0465 1.67 0.0953 .
## LastAuthorFemale1 -0.0642 0.0500 -1.28 0.2002
## Year1997 -0.2332 0.1672 -1.39 0.1642
## Year1998 -0.1561 0.1542 -1.01 0.3122
## Year1999 -0.2247 0.1285 -1.75 0.0815 .
## Year2000 -0.2827 0.1321 -2.14 0.0333 *
## Year2001 -0.0592 0.1388 -0.43 0.6702
## Year2002 -0.1818 0.1198 -1.52 0.1303
## Year2003 -0.5376 0.1646 -3.27 0.0012 **
## Year2004 -0.2771 0.1335 -2.07 0.0389 *
## Year2005 -0.2885 0.1290 -2.24 0.0261 *
```

```

## Year2006          -0.1029      0.1190   -0.86   0.3880
## Year2007          -0.1488      0.1102   -1.35   0.1781
## Year2008          -0.3247      0.1154   -2.81   0.0052 **
## Year2009          -0.3351      0.1302   -2.57   0.0106 *
## Year2010          -0.1233      0.1334   -0.92   0.3561
## Year2011          -0.1812      0.1275   -1.42   0.1562
## Year2012          -0.2015      0.1405   -1.43   0.1526
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.356
## Multiple R-squared:  0.111, Adjusted R-squared:  0.0555
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 277 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.192  0.857  0.955  0.893  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      3.26e-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 1.289 1      1.135
## Year              1.289 16      1.008

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.03234 -0.20631 -0.00244 0.22955 1.24642
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1191 0.0981 11.40 <2e-16 ***
## FirstAuthorFemale1 0.0725 0.0465 1.56 0.120
## Year1997 -0.1976 0.1607 -1.23 0.220
## Year1998 -0.1433 0.1566 -0.91 0.361
## Year1999 -0.2102 0.1314 -1.60 0.111
## Year2000 -0.2747 0.1356 -2.03 0.044 *
## Year2001 -0.0484 0.1410 -0.34 0.732
## Year2002 -0.1565 0.1183 -1.32 0.187
## Year2003 -0.5054 0.1619 -3.12 0.002 **
## Year2004 -0.2467 0.1324 -1.86 0.063 .
## Year2005 -0.2676 0.1296 -2.06 0.040 *
## Year2006 -0.0868 0.1203 -0.72 0.471
```

```

## Year2007          -0.1283      0.1114   -1.15    0.250
## Year2008          -0.3034      0.1154   -2.63    0.009 **
## Year2009          -0.3110      0.1283   -2.42    0.016 *
## Year2010          -0.0928      0.1332   -0.70    0.486
## Year2011          -0.1532      0.1279   -1.20    0.232
## Year2012          -0.1746      0.1408   -1.24    0.216
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.357
## Multiple R-squared:  0.106, Adjusted R-squared:  0.0539
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 40 weights are ~= 1. The remaining 267 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.197  0.848  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      3.26e-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 1.286 1      1.134
## Year      1.286 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.08454 -0.20146  0.00805  0.23032  1.20449

```

```

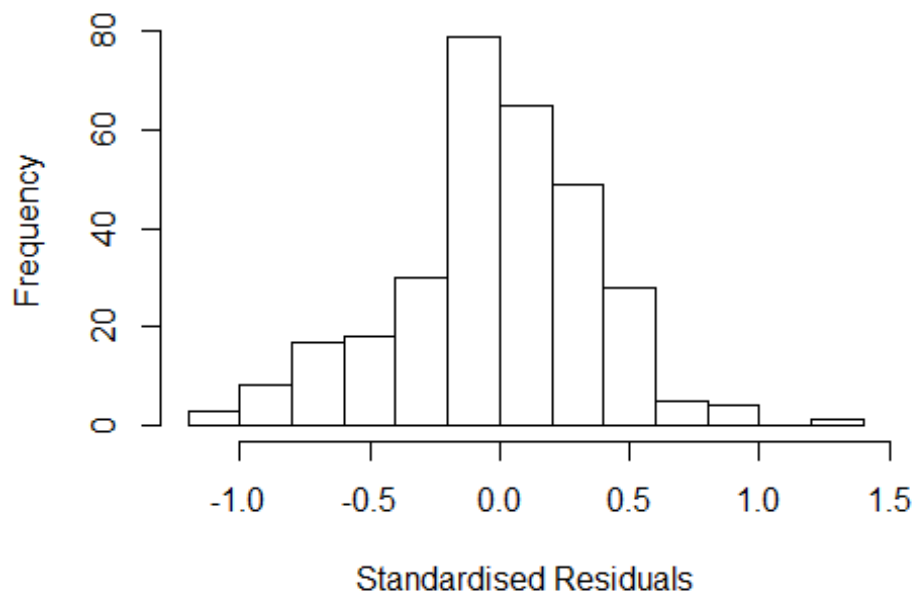
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1892    0.0990   12.01  <2e-16 ***
## LastAuthorFemale1 -0.0548    0.0507   -1.08   0.2807
## Year1997          -0.2257    0.1668   -1.35   0.1770
## Year1998          -0.1493    0.1524   -0.98   0.3280
## Year1999          -0.2216    0.1248   -1.78   0.0768 .
## Year2000          -0.2902    0.1303   -2.23   0.0267 *
## Year2001          -0.0691    0.1383   -0.50   0.6177
## Year2002          -0.1706    0.1181   -1.44   0.1496
## Year2003          -0.5515    0.1687   -3.27   0.0012 **
## Year2004          -0.2902    0.1355   -2.14   0.0330 *
## Year2005          -0.2690    0.1320   -2.04   0.0424 *
## Year2006          -0.1111    0.1194   -0.93   0.3528
## Year2007          -0.1429    0.1113   -1.28   0.2001
## Year2008          -0.3303    0.1171   -2.82   0.0051 **
## Year2009          -0.3181    0.1302   -2.44   0.0152 *
## Year2010          -0.1046    0.1340   -0.78   0.4356
## Year2011          -0.1868    0.1270   -1.47   0.1423
## Year2012          -0.2115    0.1405   -1.51   0.1332
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.358
## Multiple R-squared:  0.102, Adjusted R-squared:  0.0492
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 275 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.235  0.858  0.953  0.893  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      3.26e-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: 307"
## [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 2009 2010 2011 2012
##    1    1    4    3    1    5    1    1    1    1    3    1    3    2    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2009 2010 2011 2012
##    0    1    4    3    0    1    1    1    1    1    2    1    3    2    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2009 2010 2011 2012
##    0    1    3    3    0    1    1    1    1    1    2    1    3    2    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 6.0822019955734"
## [1] "Male first author team size 2018 geometric mean: 1"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

Residuals from last author

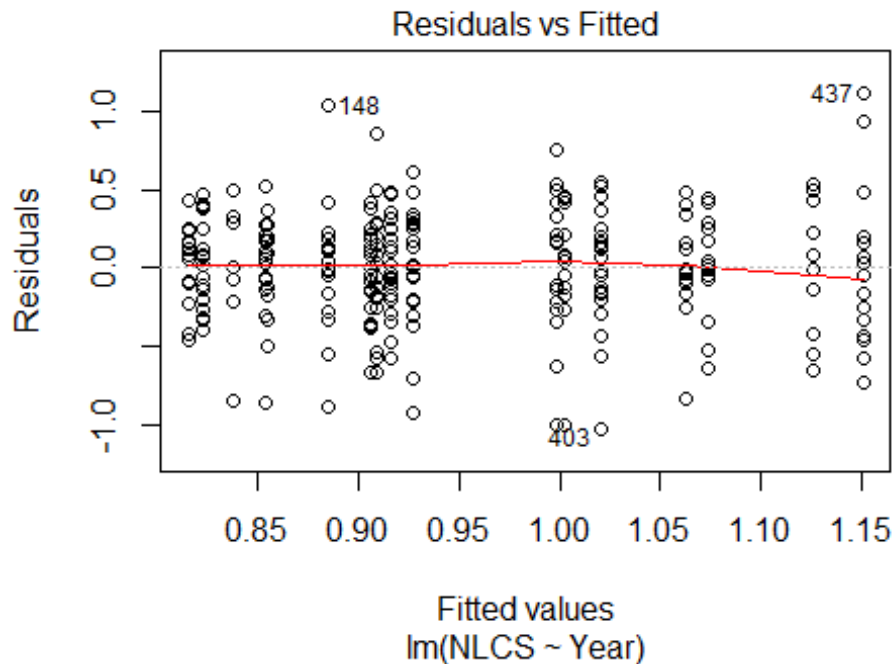


```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
```

```

## W = 3, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 3.87298334620742"
## [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: 21"
## [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
## 18 14 13 17 14 13 17 18 20 15 26 22 27 25 24
## 2011 2012
## 31 22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 15 11 7 13 13 9 14 16 17 13 24 19 20 21 17
## 2011 2012
## 22 15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 14 9 6 12 10 7 12 11 15 9 24 17 18 21 17
## 2011 2012
## 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 = 18, df = 16, p-value = 0.3

```

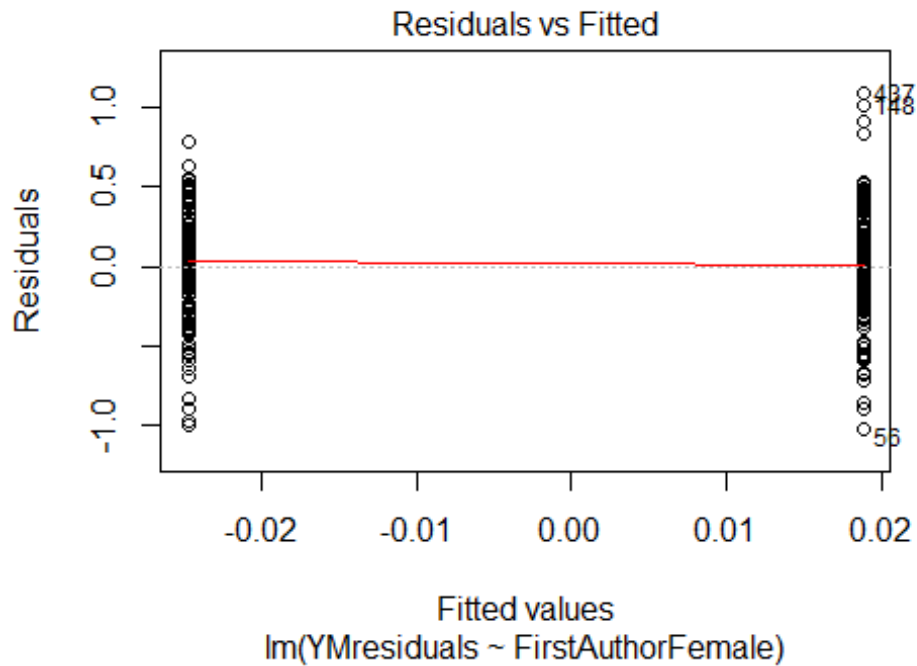
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.12, df = 1, p-value = 0.7

## [1] "Female first author team size 2018 geometric mean: 3.34676535676435"
## [1] "Male first author team size 2018 geometric mean: 1.94558479184417"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

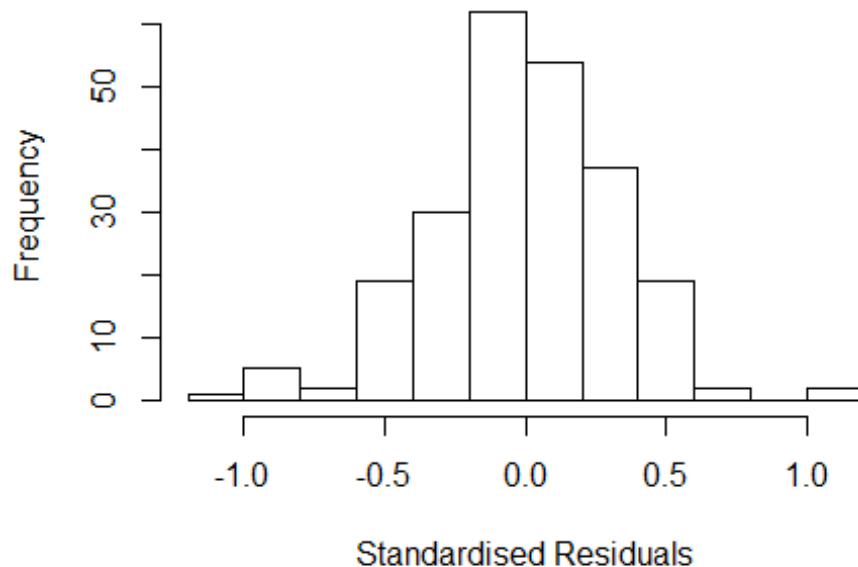
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.04
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.06135971520501"
## [1] "Male last author team size 2018 geometric mean: 2.48723569707321"

## 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 = 74, 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 1.724 1          1.313
## LastAuthorFemale  1.235 1          1.111
## UniqueAuthors    2.718 4          1.133
## Year              4.202 16         1.046
```

Residuals from first and last author and team size



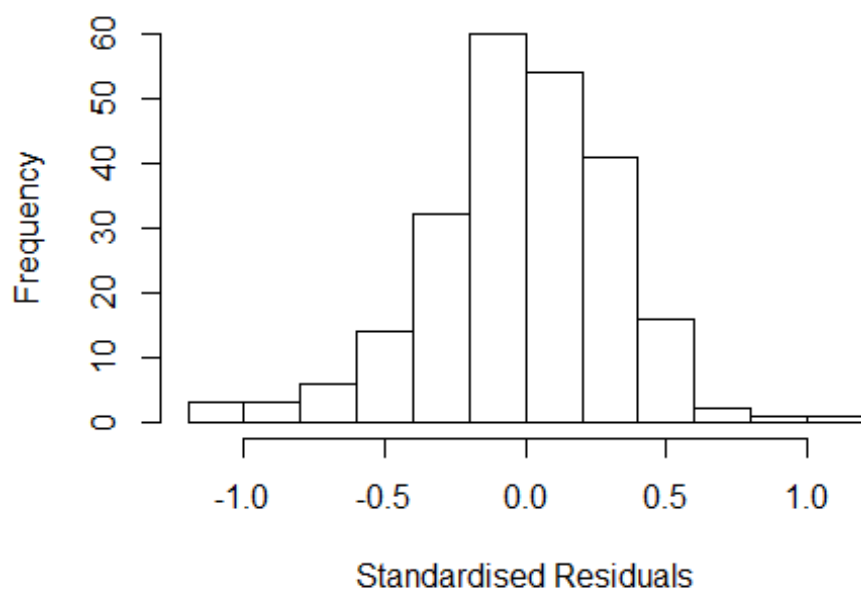
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.03348 -0.19219 -0.00755 0.20541 1.04376
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.76299 0.09251 8.25 1.8e-14 ***
## FirstAuthorFemale1 -0.03585 0.05108 -0.70 0.4836
## LastAuthorFemale1 -0.01656 0.05313 -0.31 0.7556
## UniqueAuthors2 0.05618 0.06226 0.90 0.3679
## UniqueAuthors3 0.17856 0.06597 2.71 0.0074 **
## UniqueAuthors4 0.12462 0.07295 1.71 0.0891 .
## UniqueAuthors5 0.26318 0.09410 2.80 0.0056 **
## Year1997 0.42512 0.18789 2.26 0.0247 *
## Year1998 0.12645 0.20390 0.62 0.5358
## Year1999 0.27049 0.12170 2.22 0.0273 *
```

```

## Year2000          0.19175    0.14547    1.32    0.1889
## Year2001          0.02902    0.13939    0.21    0.8353
## Year2002          0.27602    0.10448    2.64    0.0089 **
## Year2003          0.11920    0.13954    0.85    0.3939
## Year2004          0.13946    0.11843    1.18    0.2403
## Year2005          0.12927    0.10714    1.21    0.2290
## Year2006          0.05407    0.10804    0.50    0.6173
## Year2007          0.00138    0.10837    0.01    0.9898
## Year2008          0.13303    0.11439    1.16    0.2462
## Year2009          0.09927    0.10345    0.96    0.3384
## Year2010          0.17339    0.14667    1.18    0.2385
## Year2011          0.16647    0.11435    1.46    0.1469
## Year2012          0.17018    0.14820    1.15    0.2522
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.312
## Multiple R-squared:  0.122, Adjusted R-squared:  0.0295
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 211 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.239  0.872  0.952  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      4.29e-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 1.551 1      1.246
## LastAuthorFemale  1.170 1      1.082
## Year              1.795 16      1.018

```

Residuals from first and last author



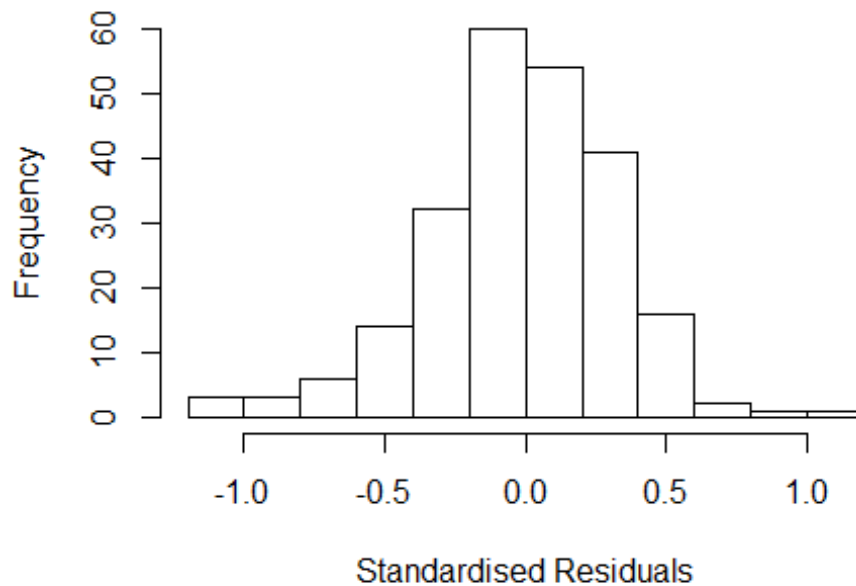
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.07116 -0.19873 -0.00312  0.21221  1.04151
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.80580    0.08820   9.14  <2e-16 ***
## FirstAuthorFemale1 0.01025    0.04934   0.21   0.836
## LastAuthorFemale1 0.00226    0.05476   0.04   0.967
## Year1997         0.40645    0.17244   2.36   0.019 *
## Year1998         0.10739    0.20824   0.52   0.607
## Year1999         0.26537    0.12225   2.17   0.031 *
## Year2000         0.24064    0.14440   1.67   0.097 .
## Year2001         0.02900    0.14028   0.21   0.836
## Year2002         0.27844    0.09819   2.84   0.005 **
## Year2003         0.19233    0.13066   1.47   0.143
## Year2004         0.16799    0.11491   1.46   0.145
## Year2005         0.18382    0.10055   1.83   0.069 .
```

```

## Year2006          0.10811      0.09898      1.09      0.276
## Year2007          0.04444      0.10753      0.41      0.680
## Year2008          0.16491      0.10994      1.50      0.135
## Year2009          0.10493      0.10138      1.03      0.302
## Year2010          0.21721      0.14234      1.53      0.128
## Year2011          0.21693      0.11298      1.92      0.056 .
## Year2012          0.23543      0.16694      1.41      0.160
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.324
## Multiple R-squared:  0.0767, Adjusted R-squared:  -0.000999
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 217 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##   0.253  0.887  0.955  0.905  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      4.29e-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 1.547 1      1.244
## Year              1.547 16      1.014

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.07217 -0.19947 -0.00351  0.21186  1.04433
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.8065     0.0865   9.33  <2e-16 ***
## FirstAuthorFemale1 0.0106     0.0493   0.22   0.8296
## Year1997          0.4070     0.1724   2.36   0.0192 *
## Year1998          0.1084     0.2088   0.52   0.6042
## Year1999          0.2656     0.1214   2.19   0.0297 *
## Year2000          0.2410     0.1448   1.66   0.0976 .
## Year2001          0.0283     0.1393   0.20   0.8391
## Year2002          0.2777     0.0974   2.85   0.0048 **
## Year2003          0.1920     0.1298   1.48   0.1405
## Year2004          0.1674     0.1144   1.46   0.1449
## Year2005          0.1837     0.1006   1.83   0.0692 .
## Year2006          0.1078     0.0991   1.09   0.2780
```

```

## Year2007          0.0440      0.1068      0.41      0.6808
## Year2008          0.1647      0.1093      1.51      0.1332
## Year2009          0.1046      0.1009      1.04      0.3014
## Year2010          0.2172      0.1431      1.52      0.1306
## Year2011          0.2169      0.1130      1.92      0.0561 .
## Year2012          0.2341      0.1657      1.41      0.1591
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.0771, Adjusted R-squared:  0.0041
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 217 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.246  0.885   0.955   0.904   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      4.29e-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 1.161 1      1.078
## 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.07314 -0.19877 -0.00477  0.20991  1.03366

```



```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.81108    0.08259   9.82  <2e-16 ***
## LastAuthorFemale1 0.00315    0.05463   0.06  0.9541
## Year1997        0.40347    0.17077   2.36  0.0190 *
## Year1998        0.10800    0.20776   0.52  0.6037
## Year1999        0.26205    0.11959   2.19  0.0295 *
## Year2000        0.23894    0.14377   1.66  0.0980 .
## Year2001        0.02689    0.13849   0.19  0.8462
## Year2002        0.27671    0.09785   2.83  0.0051 **
## Year2003        0.18792    0.12604   1.49  0.1374
## Year2004        0.16573    0.11487   1.44  0.1505
## Year2005        0.18048    0.09826   1.84  0.0676 .
## Year2006        0.10569    0.09798   1.08  0.2819
## Year2007        0.04243    0.10677   0.40  0.6914
## Year2008        0.16503    0.11040   1.49  0.1364
## Year2009        0.10668    0.10170   1.05  0.2954
## Year2010        0.21520    0.13929   1.54  0.1238
## Year2011        0.21893    0.11259   1.94  0.0531 .
## Year2012        0.23711    0.16596   1.43  0.1545
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.323
## Multiple R-squared:  0.0765, Adjusted R-squared:  0.00349
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 215 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.248  0.883   0.953   0.903   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      4.29e-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: 233"
## [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
##    4    4    2    2    1    2    2    4    4    4    4    3    3    6    2
## 2011 2012
##    6    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    3    1    2    1    0    1    4    3    4    4    2    3    5    2
## 2011 2012
##    6    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    3    1    1    1    0    1    3    3    3    4    2    3    5    2
## 2011 2012
##    5    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.64575131106459"
## [1] "Male first author team size 2018 geometric mean: 2.37841423000544"

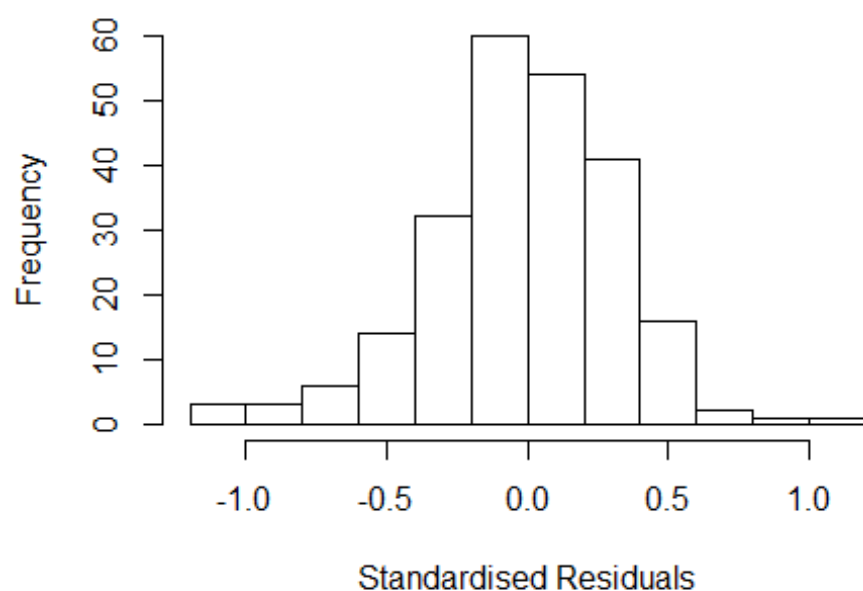
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 2.9515463231891"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

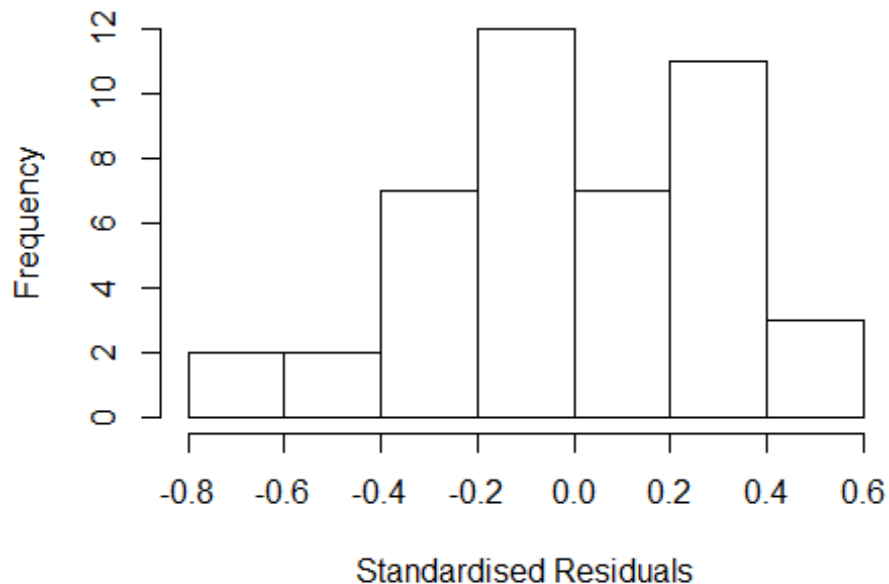
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, 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 -1.701e+14  1      NaN
## LastAuthorFemale  -1.069e+15  1      NaN
## UniqueAuthors    -2.719e+45  4      NaN
## Year              -2.170e+46 15      NaN
```

Residuals from first and last author and team size



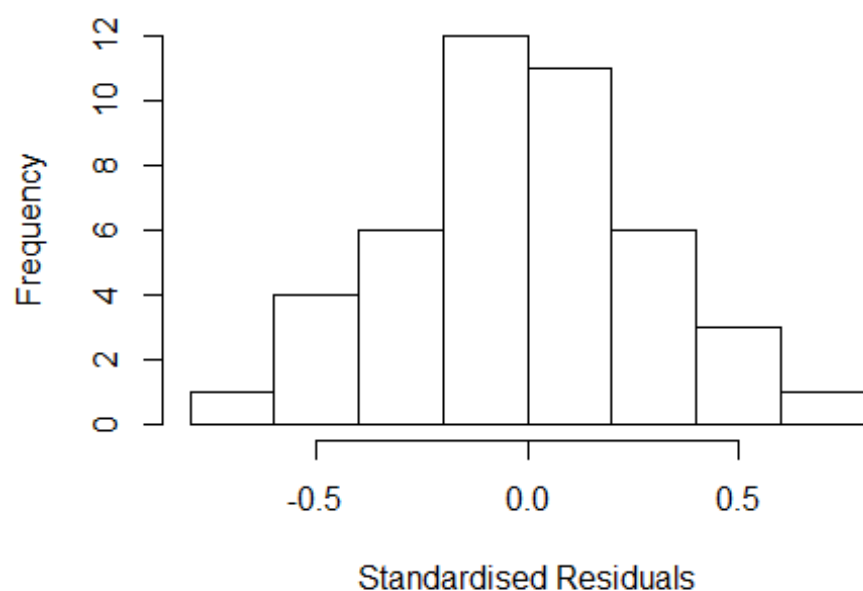
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -6.64e-01 -1.75e-01 2.36e-16 2.21e-01 5.79e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.76223 0.28804 2.65 0.015 *
## FirstAuthorFemale1 0.06586 0.11213 0.59 0.563
## LastAuthorFemale1 0.00582 0.15436 0.04 0.970
## UniqueAuthors2 -0.10166 0.29744 -0.34 0.736
## UniqueAuthors3 -0.12250 0.25559 -0.48 0.636
## UniqueAuthors4 0.02143 0.24883 0.09 0.932
## UniqueAuthors5 -0.12616 0.28054 -0.45 0.657
## Year1997 0.32057 0.50439 0.64 0.532
## Year1998 -0.41923 0.28804 -1.46 0.160
## Year1999 0.80827 0.28803 2.81 0.010 *
```

```

## Year2000          0.66025    0.30172    2.19    0.040 *
## Year2002          0.20042    0.28574    0.70    0.490
## Year2003          0.49675    0.30049    1.65    0.113
## Year2004          0.33663    0.35210    0.96    0.349
## Year2005          0.28002    0.22682    1.23    0.230
## Year2006          0.42911    0.22572    1.90    0.070 .
## Year2007         -0.08000    0.39170   -0.20    0.840
## Year2008         -0.04965    0.35622   -0.14    0.890
## Year2009          0.14669    0.35433    0.41    0.683
## Year2010          0.24310    0.34288    0.71    0.486
## Year2011          0.23655    0.25151    0.94    0.357
## Year2012          0.23766    0.30735    0.77    0.448
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.344, Adjusted R-squared:  -0.283
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 34 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.721  0.907   0.945   0.930   0.979   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      2.27e-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 -8.904e+13  1      NaN
## LastAuthorFemale  -1.231e+14  1      NaN
## Year              -3.466e+14 15      NaN

```

Residuals from first and last author



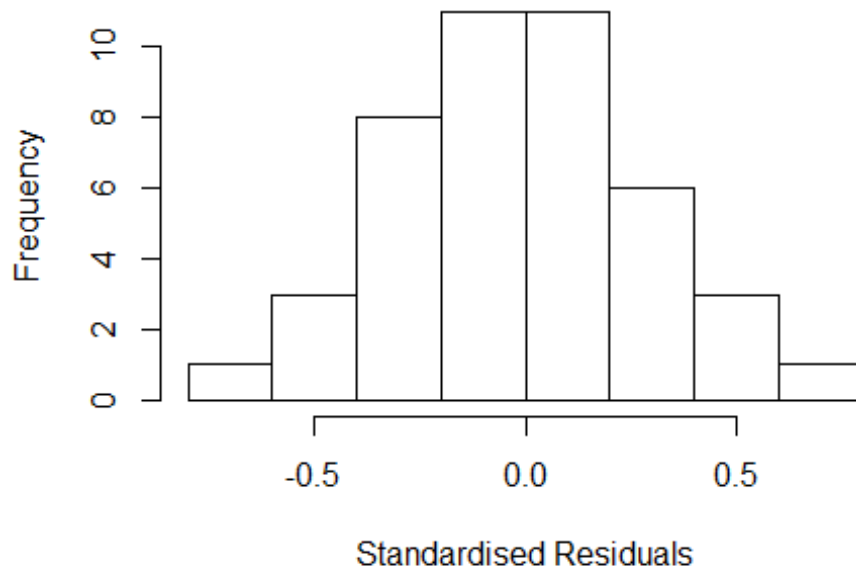
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -6.40e-01 -1.97e-01 6.25e-16 1.81e-01 6.77e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7170 0.2238 3.20 0.0036 **
## FirstAuthorFemale1 0.0400 0.0917 0.44 0.6660
## LastAuthorFemale1 0.0643 0.1223 0.53 0.6034
## Year1997 0.2682 0.4555 0.59 0.5611
## Year1998 -0.3740 0.2238 -1.67 0.1067
## Year1999 0.7310 0.2238 3.27 0.0031 **
## Year2000 0.5466 0.1709 3.20 0.0036 **
## Year2002 0.1440 0.2238 0.64 0.5257
## Year2003 0.4297 0.2565 1.67 0.1059
## Year2004 0.2699 0.2538 1.06 0.2975
## Year2005 0.2806 0.1983 1.42 0.1689
## Year2006 0.3950 0.1856 2.13 0.0430 *
```

```

## Year2007          -0.1481      0.3373   -0.44   0.6643
## Year2008          -0.0943      0.3158   -0.30   0.7675
## Year2009           0.1156      0.2785    0.42   0.6815
## Year2010           0.1306      0.2844    0.46   0.6499
## Year2011           0.2043      0.1910    1.07   0.2946
## Year2012           0.1747      0.2772    0.63   0.5340
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.322, Adjusted R-squared:  -0.121
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 36 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.799  0.940  0.969  0.954  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      2.27e-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 7.256e+14 1      2.694e+07
## Year              7.256e+14 15      3.129e+00

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -6.40e-01 -2.36e-01  5.83e-16  1.91e-01  6.77e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.7452     0.1991   3.74 0.00087 ***
## FirstAuthorFemale1 0.0525     0.0974   0.54 0.59453
## Year1997        0.2398     0.4493   0.53 0.59795
## Year1998       -0.4022     0.1991  -2.02 0.05340 .
## Year1999        0.7028     0.1991   3.53 0.00151 **
## Year2000        0.5703     0.1713   3.33 0.00252 **
## Year2002        0.1158     0.1991   0.58 0.56567
## Year2003        0.3978     0.2358   1.69 0.10314
## Year2004        0.2757     0.2450   1.13 0.27032
## Year2005        0.2696     0.1860   1.45 0.15860
## Year2006        0.4063     0.1899   2.14 0.04156 *
## Year2007       -0.1824     0.3162  -0.58 0.56877
```

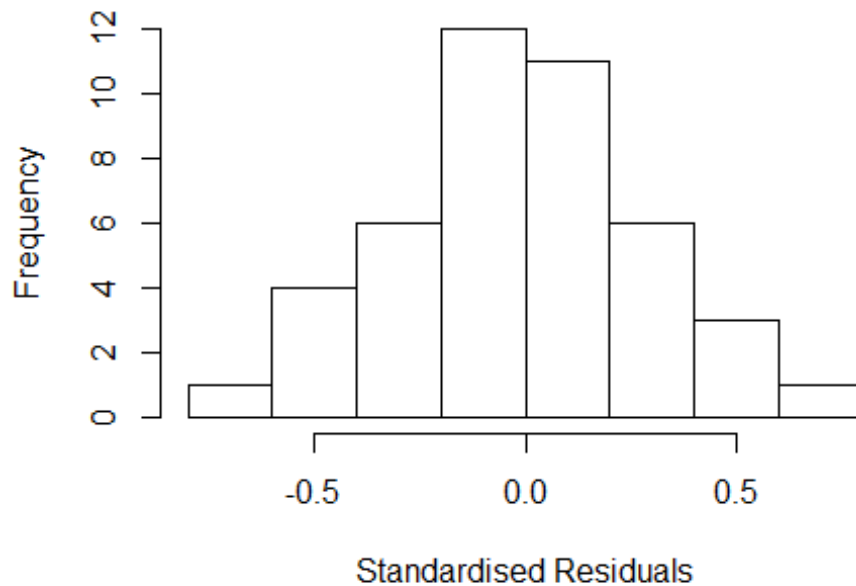


```

## Year2008          -0.1131      0.3259   -0.35   0.73121
## Year2009          0.0849      0.2520    0.34   0.73870
## Year2010          0.1606      0.2924    0.55   0.58741
## Year2011          0.2069      0.1930    1.07   0.29337
## Year2012          0.1533      0.2431    0.63   0.53366
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.319, Adjusted R-squared:  -0.0844
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 37 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.792  0.936  0.972  0.953  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      2.27e-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 -2.351e+15  1      NaN
## Year             -2.351e+15 15      NaN

```

Residuals from last author



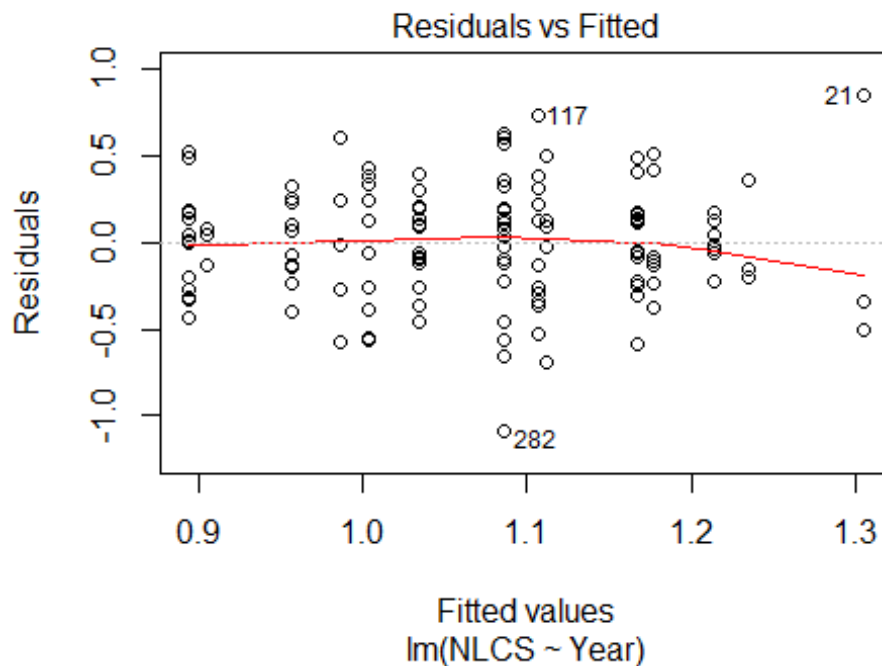
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -6.39e-01 -1.90e-01 9.44e-16 1.88e-01 6.78e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7328 0.2117 3.46 0.0018 **
## LastAuthorFemale1 0.0748 0.1282 0.58 0.5645
## Year1997 0.2516 0.4610 0.55 0.5897
## Year1998 -0.3898 0.2117 -1.84 0.0767 .
## Year1999 0.7152 0.2117 3.38 0.0022 **
## Year2000 0.5604 0.1736 3.23 0.0033 **
## Year2002 0.1282 0.2117 0.61 0.5498
## Year2003 0.4266 0.2472 1.73 0.0958 .
## Year2004 0.2742 0.2526 1.09 0.2874
## Year2005 0.2748 0.1974 1.39 0.1751
## Year2006 0.4012 0.1887 2.13 0.0428 *
## Year2007 -0.1438 0.3578 -0.40 0.6910
```

```

## Year2008          -0.0905      0.3242    -0.28    0.7823
## Year2009          0.1104      0.2811     0.39    0.6975
## Year2010          0.1244      0.2711     0.46    0.6499
## Year2011          0.2062      0.1947     1.06    0.2988
## Year2012          0.1967      0.2844     0.69    0.4949
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.322, Adjusted R-squared:  -0.0802
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.781  0.937  0.972  0.952  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      2.27e-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: 44"
## [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
##    6   12   11    9    9    5   12    7   12   20   18   10   16   19   21
## 2011 2012
##   17   22
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    7    3    6    3    5    3    8   12   16    5    7   13   14
## 2011 2012
##   10   13

```

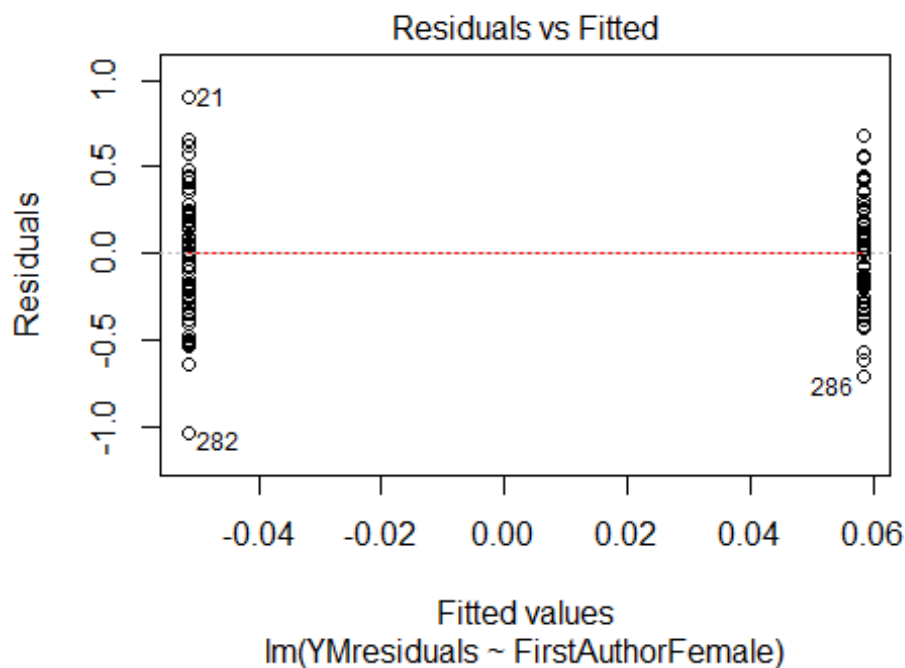
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    7    3    5    3    4    2    7   10   13    4    7   12   12
## 2011 2012
##    8   12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 30, df = 16, p-value = 0.02
```



```
##
## 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: 4.45007854298396"
## [1] "Male first author team size 2018 geometric mean: 5.72685440944837"
## 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 = 9, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.55655882007785"
## [1] "Male last author team size 2018 geometric mean: 5.82779555555977"

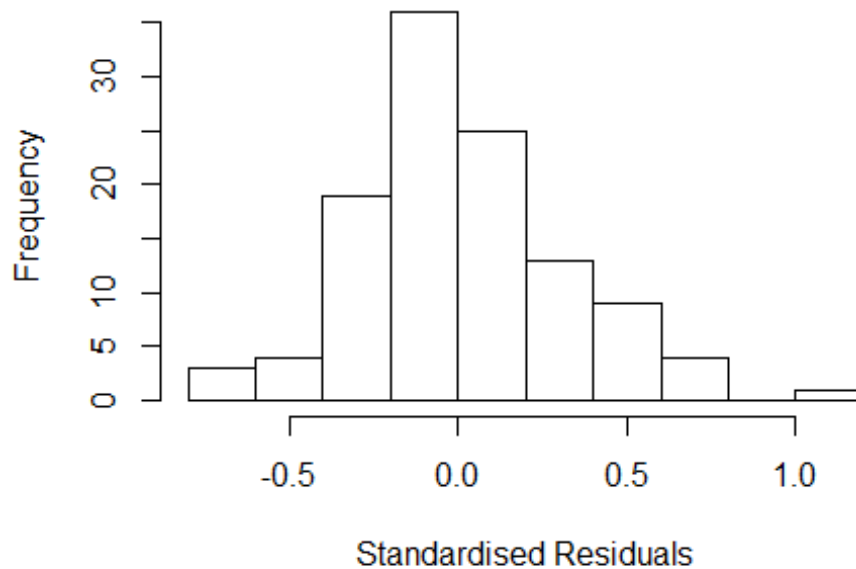
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 6, 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	2.038	1	1.428
LastAuthorFemale	4.423	1	2.103
UniqueAuthors	24.815	4	1.494
Year	110.237	16	1.158

Residuals from first and last author and team size



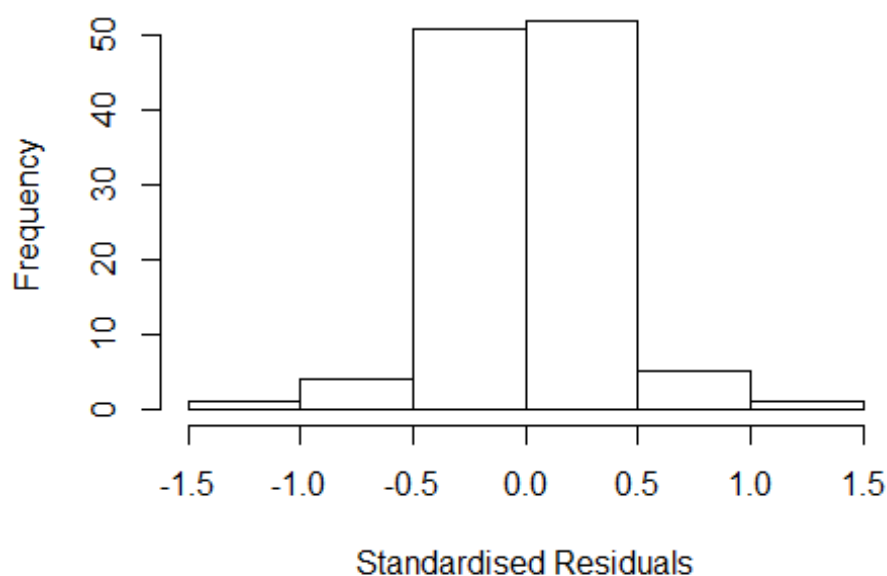
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7305 -0.1747 -0.0288 0.1899 1.0136
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.48774 0.51364 0.95 0.34
## FirstAuthorFemale1 0.06856 0.06695 1.02 0.31
## LastAuthorFemale1 0.07707 0.11282 0.68 0.50
## UniqueAuthors2 0.47144 0.36491 1.29 0.20
## UniqueAuthors3 0.49265 0.40612 1.21 0.23
## UniqueAuthors4 0.58098 0.36109 1.61 0.11
## UniqueAuthors5 0.53787 0.37141 1.45 0.15
## Year1997 -0.00435 0.61576 -0.01 0.99
## Year1998 0.19896 0.36272 0.55 0.58
## Year1999 0.18069 0.39474 0.46 0.65
```

```

## Year2000      0.04604      0.35874      0.13      0.90
## Year2001     -0.15732      0.36638     -0.43      0.67
## Year2002      0.09179      0.38960      0.24      0.81
## Year2003      0.24372      0.36119      0.67      0.50
## Year2004      0.03664      0.35714      0.10      0.92
## Year2005      0.03606      0.38100      0.09      0.92
## Year2006      0.17720      0.36406      0.49      0.63
## Year2007      0.14511      0.39247      0.37      0.71
## Year2008      0.07631      0.38113      0.20      0.84
## Year2009     -0.18392      0.37121     -0.50      0.62
## Year2010      0.01810      0.37212      0.05      0.96
## Year2011     -0.22943      0.37520     -0.61      0.54
## Year2012      0.05472      0.42711      0.13      0.90
##
## Robust residual standard error: 0.315
## Multiple R-squared:  0.219, Adjusted R-squared:  0.03
## Convergence in 31 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 105 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.280  0.885  0.964  0.914  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      8.77e-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 1.943 1      1.394
## LastAuthorFemale  3.465 1      1.861
## Year              5.986 16      1.058

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.10804 -0.19402  0.00697  0.19189  1.14167
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0030     0.3761    2.67   0.009 **
## FirstAuthorFemale1  0.0918     0.0666    1.38   0.171
## LastAuthorFemale1  0.0772     0.0989    0.78   0.437
## Year1997         -0.0668     0.4945   -0.14   0.893
## Year1998          0.1707     0.3809    0.45   0.655
## Year1999          0.1880     0.4173    0.45   0.653
## Year2000          0.0143     0.3924    0.04   0.971
## Year2001         -0.1587     0.3849   -0.41   0.681
## Year2002          0.0900     0.4172    0.22   0.830
## Year2003          0.2321     0.4059    0.57   0.569
## Year2004          0.0182     0.3897    0.05   0.963
## Year2005          0.0532     0.4001    0.13   0.895
```

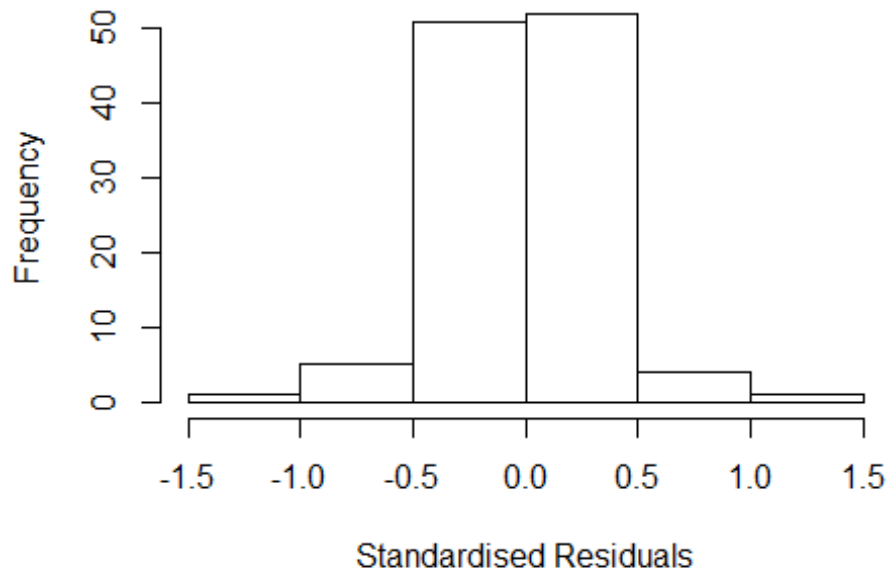


```

## Year2006          0.1283      0.3851      0.33      0.740
## Year2007          0.0975      0.4540      0.21      0.830
## Year2008          0.1024      0.3954      0.26      0.796
## Year2009         -0.1704      0.3894     -0.44      0.663
## Year2010          0.0287      0.3870      0.07      0.941
## Year2011         -0.2257      0.4000     -0.56      0.574
## Year2012          0.1050      0.4622      0.23      0.821
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.318
## Multiple R-squared:  0.154, Adjusted R-squared:  -0.00652
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 106 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.170  0.877  0.961  0.909  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      8.77e-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 1.844 1      1.358
## Year              1.844 16      1.019

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.149 -0.189 0.017 0.169 1.259
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0030 0.3721 2.70 0.0083 **
## FirstAuthorFemale1 0.0885 0.0660 1.34 0.1828
## Year1997 -0.1069 0.4097 -0.26 0.7946
## Year1998 0.1722 0.3770 0.46 0.6488
## Year1999 0.1895 0.4131 0.46 0.6474
## Year2000 0.0337 0.3858 0.09 0.9305
## Year2001 -0.1565 0.3809 -0.41 0.6820
## Year2002 0.0916 0.4136 0.22 0.8252
## Year2003 0.2337 0.4016 0.58 0.5620
## Year2004 0.0189 0.3857 0.05 0.9610
## Year2005 0.0644 0.3947 0.16 0.8708
## Year2006 0.1497 0.3789 0.40 0.6937
```

```

## Year2007          0.1281      0.4435      0.29      0.7733
## Year2008          0.1116      0.3896      0.29      0.7751
## Year2009         -0.1482      0.3837     -0.39      0.7002
## Year2010          0.0568      0.3793      0.15      0.8813
## Year2011         -0.1808      0.3857     -0.47      0.6404
## Year2012          0.1462      0.4360      0.34      0.7381
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.322
## Multiple R-squared:  0.149, Adjusted R-squared:  -0.0014
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0921 0.8800 0.9690 0.9120 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      8.77e-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 3.623 1      1.903
## Year      3.623 16      1.041

## [1] "List of 0 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
## -1.18300 -0.18807 -0.00408  0.17023  1.12008

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0030     0.3858   2.60  0.011 *
## LastAuthorFemale1  0.0630     0.0998   0.63  0.530
## Year1997         -0.0310     0.5053  -0.06  0.951
## Year1998          0.2115     0.3885   0.54  0.587
## Year1999          0.2198     0.4183   0.53  0.601
## Year2000          0.0920     0.3918   0.23  0.815
## Year2001         -0.0978     0.3896  -0.25  0.802
## Year2002          0.1199     0.4297   0.28  0.781
## Year2003          0.2780     0.4014   0.69  0.490
## Year2004          0.0644     0.3984   0.16  0.872
## Year2005          0.0848     0.4117   0.21  0.837
## Year2006          0.1792     0.3923   0.46  0.649
## Year2007          0.1672     0.4606   0.36  0.717
## Year2008          0.1436     0.4091   0.35  0.726
## Year2009         -0.1360     0.3968  -0.34  0.733
## Year2010          0.0617     0.3904   0.16  0.875
## Year2011         -0.1435     0.3962  -0.36  0.718
## Year2012          0.1800     0.4611   0.39  0.697
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.309
## Multiple R-squared:  0.141, Adjusted R-squared:  -0.0105
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~ = 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.110  0.865   0.968   0.905   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      8.77e-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: 114"
## [1] ""

```

```

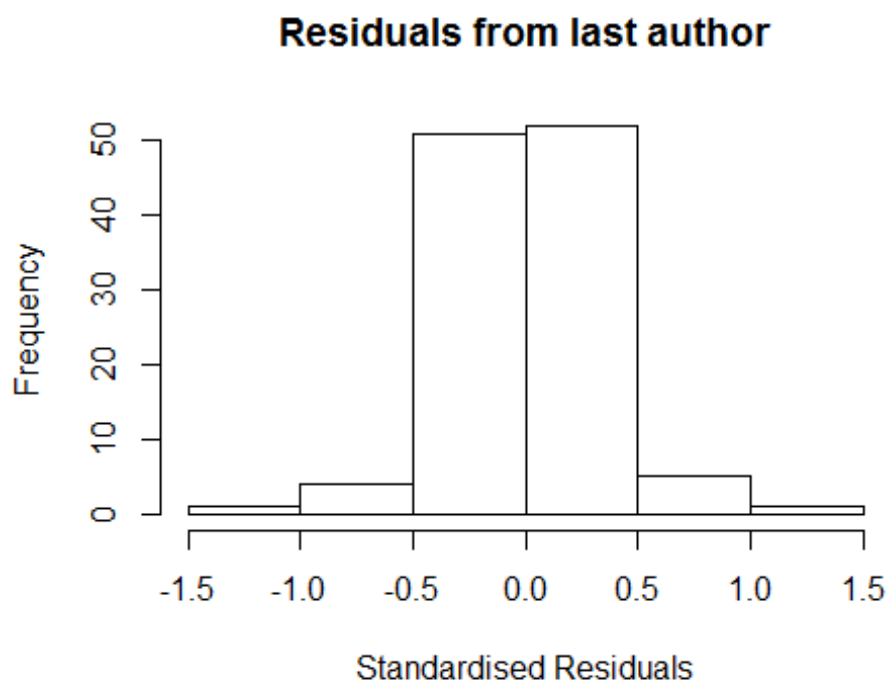
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2805"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    3    8    6    4    3    6    9   10    9   11   13   12   17   21
## 2011 2012
##   18   24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    1    7    2    4    2    6    5   10    5    8   11    7    9   13
## 2011 2012
##   11   21
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    1    7    2    3    2    6    4    8    4    7    7    6    7   11
## 2011 2012
##   11   18
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.68179283050743"
## [1] "Male first author team size 2018 geometric mean: 2.7200434229974"

## 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 = 7, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.96798967126543"
## [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

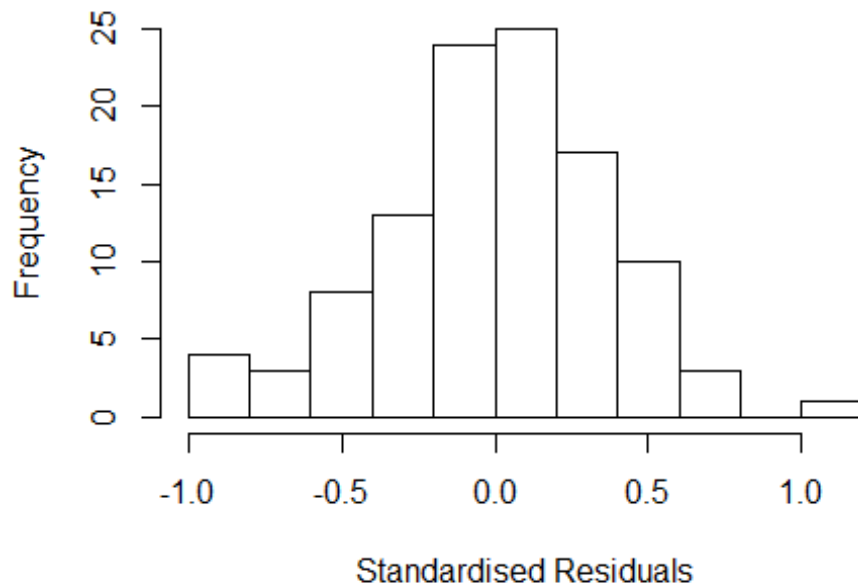
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, p-value = 0.8
## 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	6.863	1	2.620
## LastAuthorFemale	7.409	1	2.722
## UniqueAuthors	220.567	4	1.963
## Year	1840.154	16	1.265

Residuals from first and last author and team size



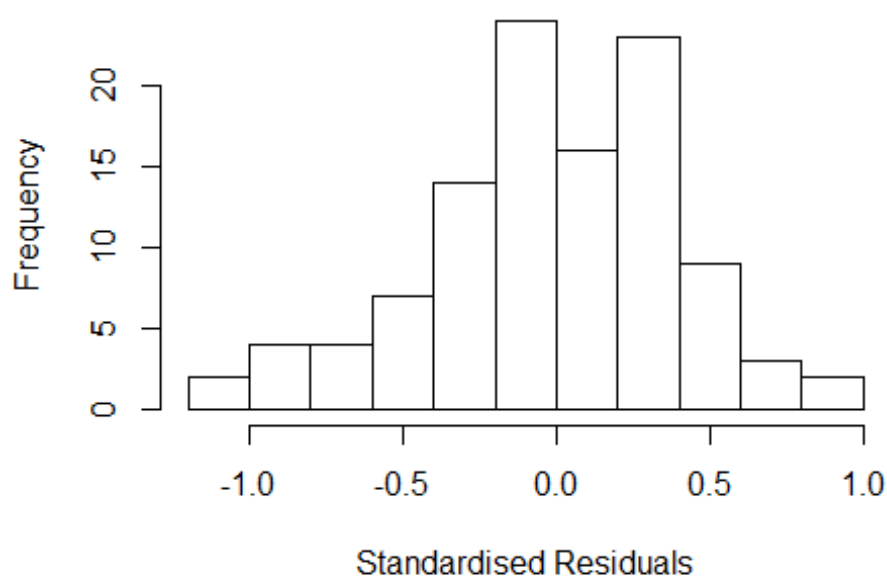
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8661 -0.2195 0.0128 0.2381 1.1070
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03373 0.11901 8.69 2.3e-13 ***
## FirstAuthorFemale1 0.01446 0.09634 0.15 0.88102
## LastAuthorFemale1 -0.23260 0.09660 -2.41 0.01821 *
## UniqueAuthors2 0.00719 0.12162 0.06 0.95302
## UniqueAuthors3 0.35944 0.13388 2.68 0.00872 **
## UniqueAuthors4 0.50424 0.18229 2.77 0.00696 **
## UniqueAuthors5 0.45560 0.14544 3.13 0.00238 **
## Year1997 0.94027 0.11901 7.90 9.0e-12 ***
## Year1998 0.01524 0.21403 0.07 0.94341
## Year1999 0.21198 0.11271 1.88 0.06342 .
```

```

## Year2000      -0.09984    0.13464   -0.74  0.46041
## Year2001      -0.40517    0.35119   -1.15  0.25185
## Year2002       0.13334    0.14495    0.92  0.36021
## Year2003      -0.30371    0.23136   -1.31  0.19280
## Year2004       0.02393    0.13005    0.18  0.85444
## Year2005      -0.14192    0.14764   -0.96  0.33918
## Year2006      -0.17477    0.17126   -1.02  0.31040
## Year2007      -0.17754    0.18719   -0.95  0.34559
## Year2008      -0.04306    0.13952   -0.31  0.75834
## Year2009      -0.00386    0.18552   -0.02  0.98344
## Year2010      -0.19969    0.27126   -0.74  0.46366
## Year2011      -0.20366    0.18861   -1.08  0.28329
## Year2012      -0.47755    0.12919   -3.70  0.00039 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.316, Adjusted R-squared:  0.139
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 95 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.409  0.896  0.952  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      9.26e-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 6.341 1          2.518
## LastAuthorFemale  5.480 1          2.341
## Year             23.723 16          1.104

```


Residuals from first and last author



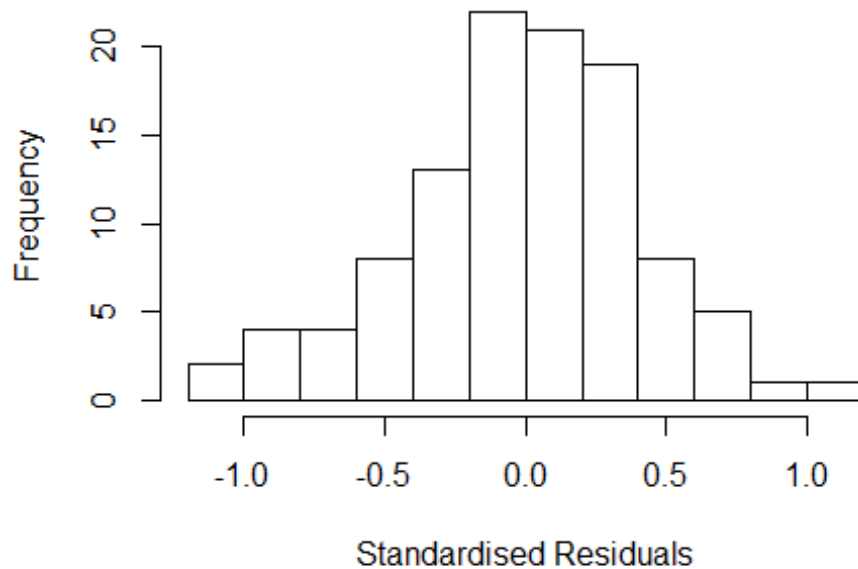
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.08169 -0.22419 -0.00111 0.27196 0.95910
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0179 0.1043 9.76 1.0e-15 ***
## FirstAuthorFemale1 0.0343 0.0978 0.35 0.726
## LastAuthorFemale1 -0.2266 0.1048 -2.16 0.033 *
## Year1997 0.9561 0.1043 9.17 1.7e-14 ***
## Year1998 0.0313 0.2104 0.15 0.882
## Year1999 0.2284 0.1091 2.09 0.039 *
## Year2000 0.0755 0.1555 0.49 0.628
## Year2001 -0.1436 0.9180 -0.16 0.876
## Year2002 0.1439 0.1227 1.17 0.244
## Year2003 0.1102 0.1949 0.57 0.573
## Year2004 0.1613 0.1290 1.25 0.214
## Year2005 0.1213 0.1266 0.96 0.341
```

```

## Year2006          -0.0815      0.1723    -0.47      0.637
## Year2007           0.1021      0.1067      0.96      0.341
## Year2008           0.1670      0.1233      1.35      0.179
## Year2009           0.1728      0.2014      0.86      0.393
## Year2010          -0.0360      0.2663     -0.14      0.893
## Year2011           0.0295      0.1708      0.17      0.863
## Year2012          -0.2197      0.1555     -1.41      0.161
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.167, Adjusted R-squared:  -0.00109
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 98 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.460  0.875   0.959   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      9.26e-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 5.557 1      2.357
## Year              5.557 16      1.055

```

Residuals from first author



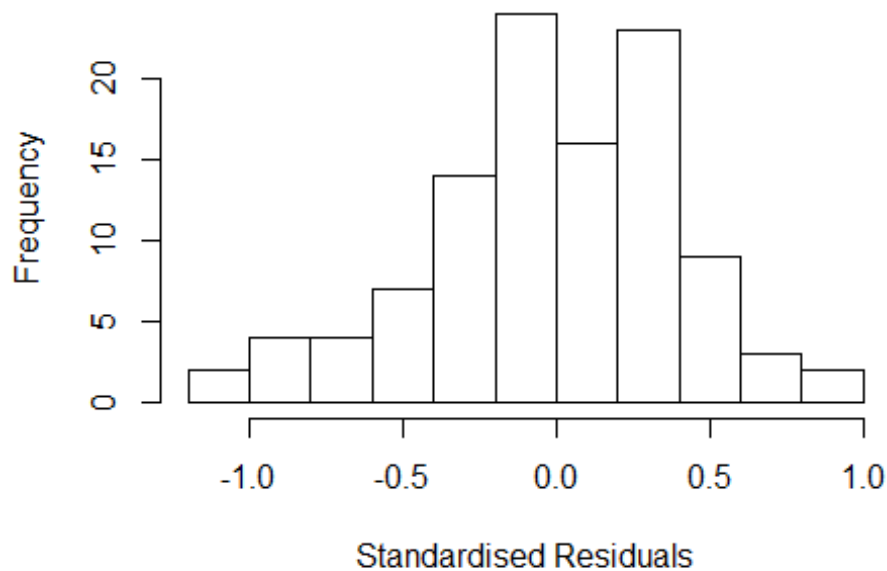
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0418 -0.2415 0.0126 0.2770 1.0084
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9621 0.1012 9.51 3.0e-15 ***
## FirstAuthorFemale1 -0.0397 0.0982 -0.40 0.69
## Year1997 1.0119 0.1012 10.00 2.9e-16 ***
## Year1998 0.0760 0.1874 0.41 0.69
## Year1999 0.1709 0.1112 1.54 0.13
## Year2000 0.0132 0.1644 0.08 0.94
## Year2001 -0.0508 0.7680 -0.07 0.95
## Year2002 0.1636 0.1099 1.49 0.14
## Year2003 0.1376 0.1592 0.86 0.39
## Year2004 0.1845 0.1546 1.19 0.24
## Year2005 0.1926 0.1461 1.32 0.19
## Year2006 -0.0144 0.1759 -0.08 0.94
```

```

## Year2007          0.1132      0.1063      1.07      0.29
## Year2008          0.1883      0.1202      1.57      0.12
## Year2009          0.1280      0.2098      0.61      0.54
## Year2010         -0.0295      0.2813     -0.10      0.92
## Year2011          0.1194      0.1581      0.76      0.45
## Year2012         -0.1701      0.1664     -1.02      0.31
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.407
## Multiple R-squared:  0.128, Adjusted R-squared:  -0.0363
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 101 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.491  0.892  0.951  0.905  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      9.26e-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 5.692 1          2.386
## Year            5.692 16          1.056

```

Residuals from last author



```
## [1] "List of 0 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
## -1.07610 -0.22819 0.00356 0.27350 0.93472
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0380 0.0878 11.82 <2e-16 ***
## LastAuthorFemale1 -0.2152 0.1018 -2.11 0.037 *
## Year1997 0.9360 0.0878 10.66 <2e-16 ***
## Year1998 0.0168 0.2129 0.08 0.937
## Year1999 0.2026 0.0811 2.50 0.014 *
## Year2000 0.0605 0.1538 0.39 0.695
## Year2001 -0.1465 0.8772 -0.17 0.868
## Year2002 0.1369 0.1233 1.11 0.270
## Year2003 0.0930 0.1901 0.49 0.626
## Year2004 0.1569 0.1341 1.17 0.245
## Year2005 0.1324 0.1259 1.05 0.296
## Year2006 -0.0940 0.1694 -0.56 0.580
```

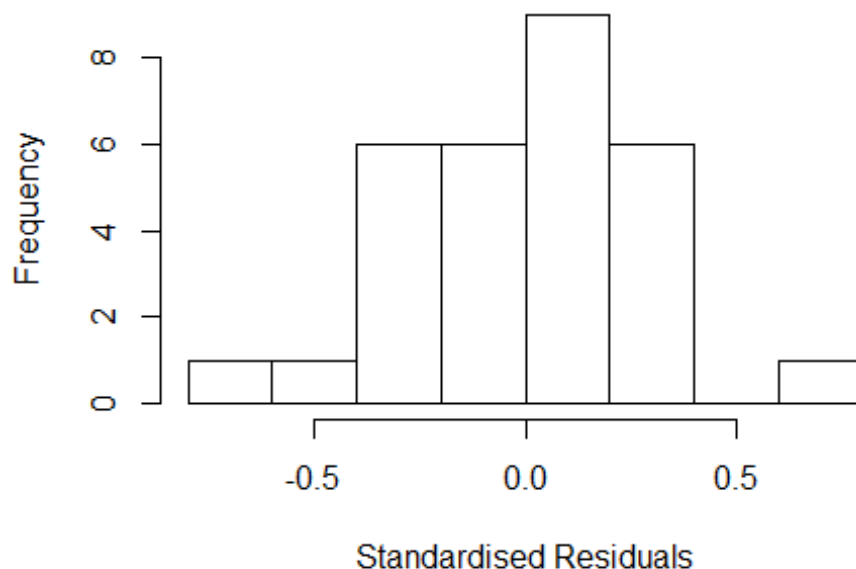
```

## Year2007          0.1021      0.1078      0.95      0.346
## Year2008          0.1599      0.1220      1.31      0.193
## Year2009          0.1619      0.2071      0.78      0.436
## Year2010         -0.0317      0.2677     -0.12      0.906
## Year2011          0.0381      0.1689      0.23      0.822
## Year2012         -0.2223      0.1571     -1.42      0.160
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.399
## Multiple R-squared:  0.168, Adjusted R-squared:  0.0112
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 97 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.448  0.870   0.958   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      9.26e-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: 108"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    3    1    3    3    1    3    2    5    1    2    3    5    2    4    3
## 2012
##    6
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    1    2    2    1    3    1    4    0    2    2    4    2    4    2
## 2012

```

```
##      5
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##      2      0      2      2      1      3      0      3      0      2      2      3      2      4      1
## 2012
##      3
## [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"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.387e+14  1      1.840e+07
## LastAuthorFemale  2.558e+14  1      1.599e+07
## Year              5.378e+15 12      4.523e+00
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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))
```

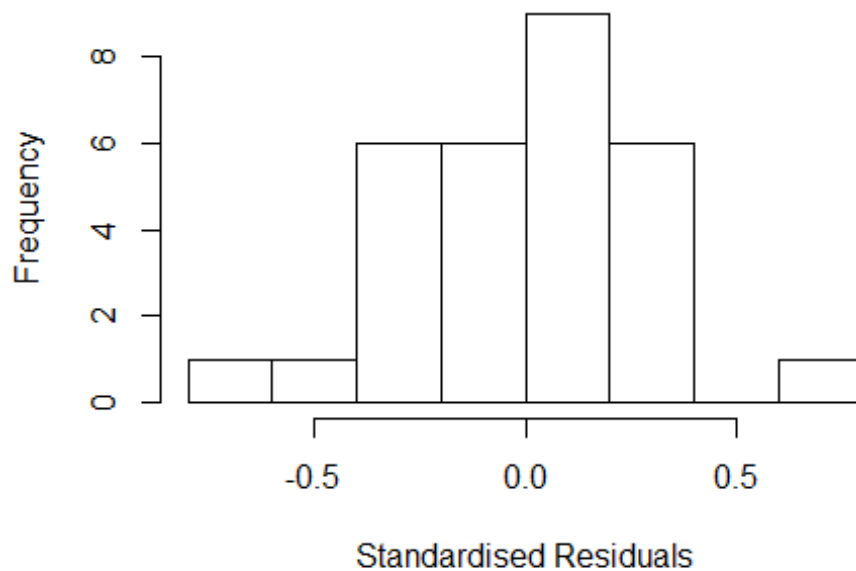
```

## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7075 -0.2076  0.0341  0.1701  0.7075
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    8.24e-01   2.75e-01   3.00   0.009 **
## FirstAuthorFemale1 1.19e-02   1.32e-01   0.09   0.929
## LastAuthorFemale1 7.74e-05   1.14e-01   0.00   0.999
## Year1999       -1.29e-01   1.00e+00  -0.13   0.899
## Year2000        2.61e-01   2.70e-01   0.97   0.349
## Year2001        3.34e-01   2.75e-01   1.21   0.244
## Year2002        1.66e-01   2.96e-01   0.56   0.582
## Year2004        1.34e-01   2.76e-01   0.49   0.633
## Year2006        4.44e-01   2.74e-01   1.62   0.127
## Year2007        2.75e-01   3.42e-01   0.80   0.435
## Year2008        1.73e-01   2.98e-01   0.58   0.571
## Year2009        1.78e-01   2.73e-01   0.65   0.525
## Year2010        4.99e-02   3.09e-01   0.16   0.874
## Year2011        3.48e-01   2.52e-01   1.38   0.189
## Year2012        7.74e-01   2.83e-01   2.73   0.015 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.349, Adjusted R-squared:  -0.258
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 24 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.796  0.949  0.972   0.956   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      3.33e-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	2.453e+15	1	4.952e+07
##	Year	2.453e+15	12	4.378e+00

Residuals from first author



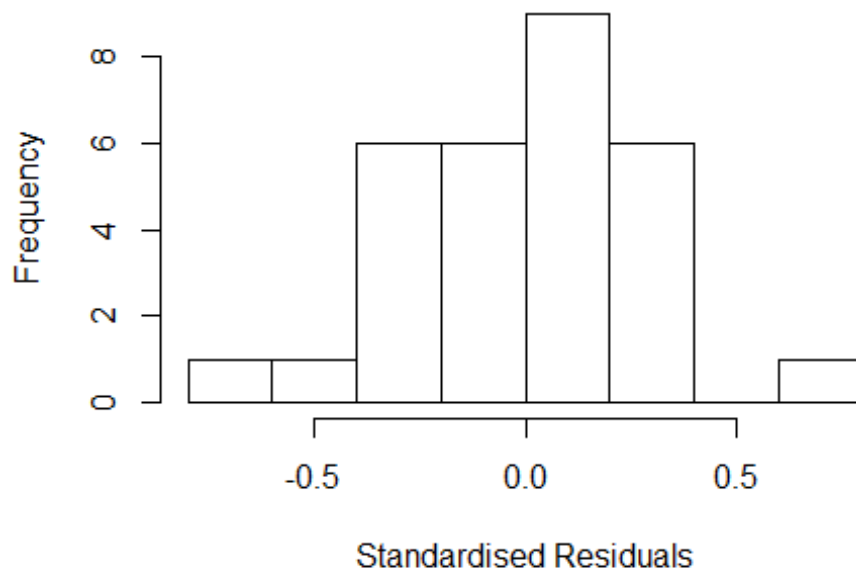
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.7075 -0.2076 0.0341 0.1701 0.7075
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8245 0.2539 3.25 0.0051 **
## FirstAuthorFemale1 0.0119 0.1325 0.09 0.9295
## Year1999 -0.1290 1.0673 -0.12 0.9053
## Year2000 0.2605 0.2457 1.06 0.3047
## Year2001 0.3335 0.2539 1.31 0.2076
## Year2002 0.1664 0.3069 0.54 0.5951
## Year2004 0.1344 0.2653 0.51 0.6194
## Year2006 0.4435 0.2755 1.61 0.1270
```

```

## Year2007          0.2745      0.3443      0.80      0.4370
## Year2008          0.1728      0.3012      0.57      0.5741
## Year2009          0.1780      0.2687      0.66      0.5172
## Year2010          0.0498      0.3032      0.16      0.8716
## Year2011          0.3475      0.2539      1.37      0.1900
## Year2012          0.7743      0.2706      2.86      0.0113 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.35,   Adjusted R-squared:  -0.178
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.784  0.947  0.970  0.954  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      3.33e-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 2.562e+14 1      1.601e+07
## Year              2.562e+14 12      3.984e+00

```

Residuals from last author



```
## [1] "List of 0 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
## -0.708 -0.207 0.034 0.174 0.707
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.830943 0.275688 3.01 0.0082 **
## LastAuthorFemale1 -0.000443 0.114501 0.00 0.9970
## Year1999 -0.123000 1.029639 -0.12 0.9064
## Year2000 0.260057 0.276451 0.94 0.3608
## Year2001 0.327057 0.275688 1.19 0.2528
## Year2002 0.167440 0.303629 0.55 0.5889
## Year2004 0.140032 0.272950 0.51 0.6149
## Year2006 0.443500 0.281845 1.57 0.1352
## Year2007 0.274500 0.343392 0.80 0.4358
## Year2008 0.178296 0.293744 0.61 0.5524
## Year2009 0.177779 0.277619 0.64 0.5310
## Year2010 0.055541 0.306309 0.18 0.8584
```

```

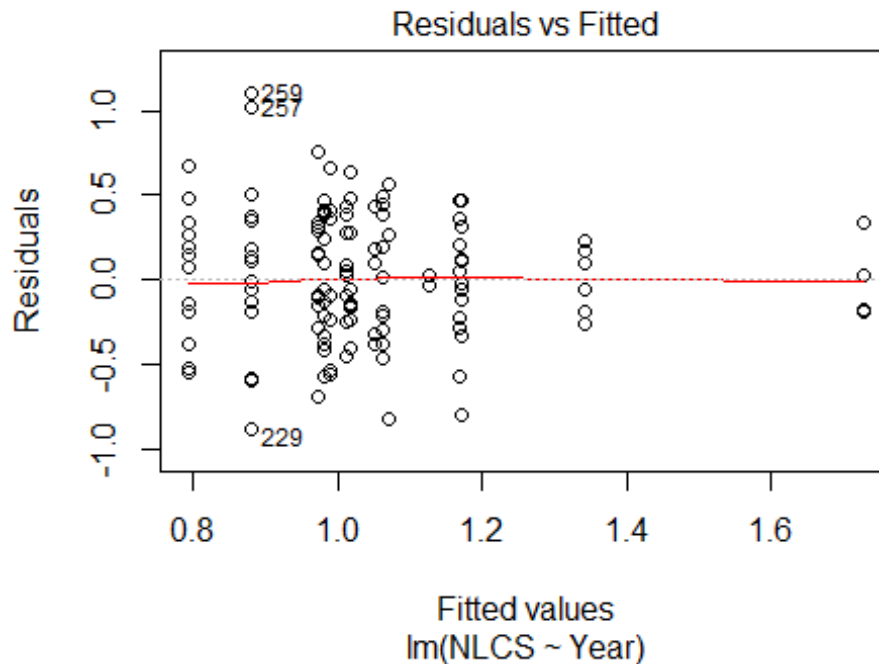
## Year2011          0.353500    0.249917    1.41    0.1764
## Year2012          0.768061    0.282598    2.72    0.0152 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.35,    Adjusted R-squared:  -0.179
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 24 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.790  0.948  0.971   0.955   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           3.33e-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: 30"
## [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]"
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2009 2010 2011 2012
##    1    1    5    2    2    2    1    3    2    4    4    8    7
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2009 2010 2011 2012
##    1    0    4    1    0    0    0    1    2    2    0    3    1
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2009 2010 2011 2012
##    1    0    4    1    0    0    0    0    2    2    0    2    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.47213595499958"
## [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: 13"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    5    9    8   14    5   10    7   12   14   12   14   11   16   15   15
## 2012
##    24
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    3    6    8    2    6    5   10   10   10   10    7   13    7   12
## 2012
##    18
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    4    3    5    8    1    6    5   10    9    9    9    5   11    6   11
## 2012
##    17
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 15, df = 15, p-value = 0.4

```



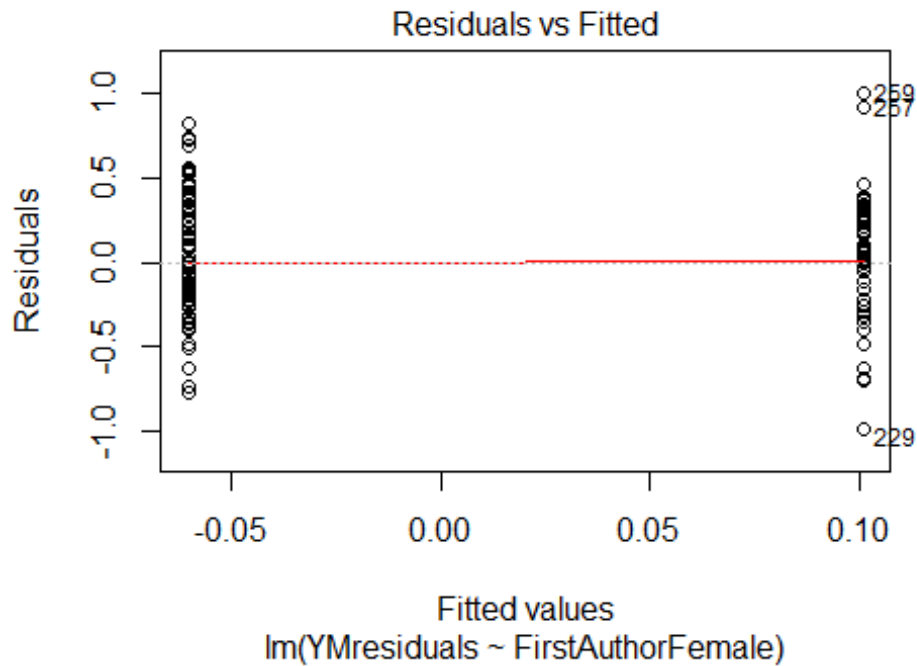
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.84, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 4.54546269184229"
## [1] "Male first author team size 2018 geometric mean: 3.68822944503737"

## 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 = 40, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.31829589694499"
## [1] "Male last author team size 2018 geometric mean: 3.6332857436926"

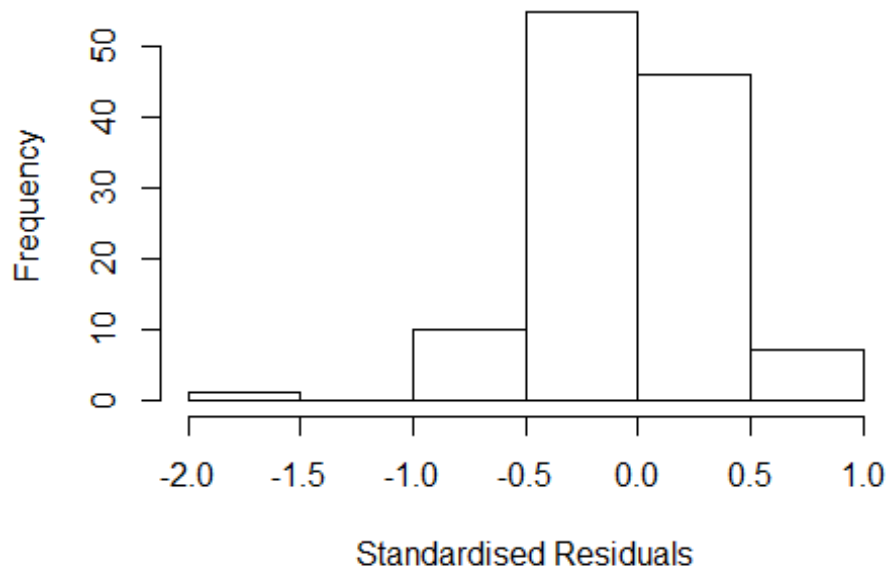
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 40, 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	3.171	1	1.781
## LastAuthorFemale	3.027	1	1.740
## UniqueAuthors	60.781	4	1.671
## Year	206.416	15	1.194

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.5112 -0.2033 -0.0316 0.2337 0.9109
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2880 0.2118 6.08 2.3e-08 ***
## FirstAuthorFemale1 0.1576 0.0841 1.88 0.06376 .
## LastAuthorFemale1 0.0851 0.0872 0.98 0.33162
## UniqueAuthors2 0.2917 0.1475 1.98 0.05081 .
## UniqueAuthors3 0.3417 0.1054 3.24 0.00163 **
## UniqueAuthors4 0.4214 0.1185 3.56 0.00058 ***
## UniqueAuthors5 0.5516 0.1305 4.23 5.4e-05 ***
## Year1998 -0.0824 0.2794 -0.29 0.76870
## Year1999 -0.4949 0.2100 -2.36 0.02048 *
## Year2000 -0.6924 0.2219 -3.12 0.00238 **
```

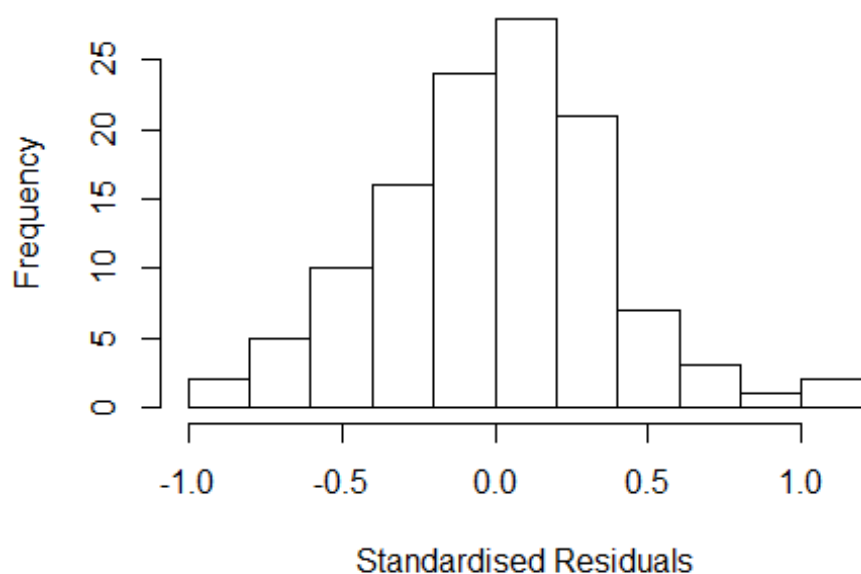


```

## Year2001          -0.6816      0.2184    -3.12   0.00238 **
## Year2002          -0.3171      0.2250    -1.41   0.16191
## Year2003          -0.7327      0.2599    -2.82   0.00583 **
## Year2004          -0.5485      0.2319    -2.37   0.02001 *
## Year2005          -0.8111      0.2057    -3.94   0.00015 ***
## Year2006          -0.7188      0.2036    -3.53   0.00063 ***
## Year2007          -0.7003      0.2399    -2.92   0.00437 **
## Year2008          -0.6964      0.2517    -2.77   0.00679 **
## Year2009          -0.8816      0.2097    -4.20   5.8e-05 ***
## Year2010          -0.9343      0.2744    -3.41   0.00096 ***
## Year2011          -0.7014      0.2203    -3.18   0.00196 **
## Year2012          -0.9241      0.2364    -3.91   0.00017 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.359
## Multiple R-squared:  0.346, Adjusted R-squared:  0.204
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 109 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0372 0.8700 0.9610 0.9090 0.9880 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          8.40e-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 2.711 1 1.647
## LastAuthorFemale 2.630 1 1.622
## Year 4.978 15 1.055

```

Residuals from first and last author



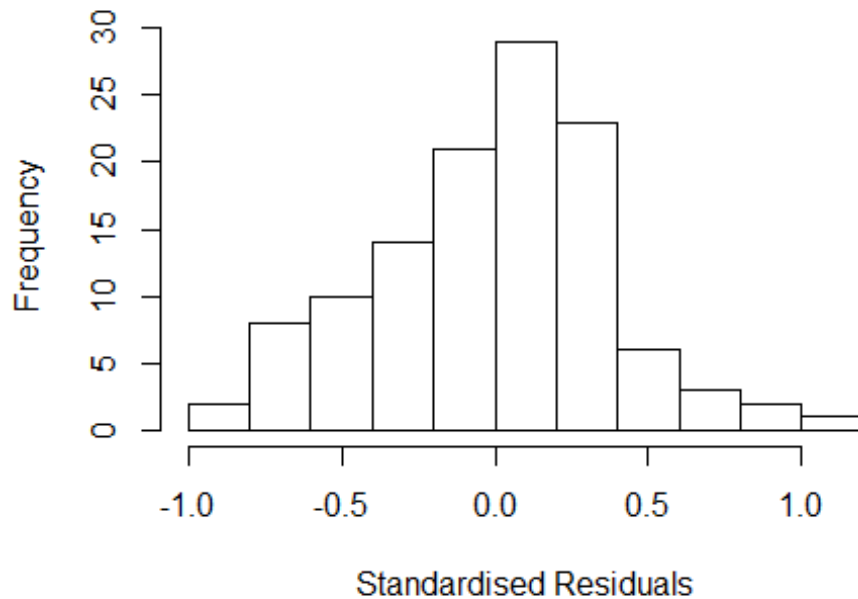
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8843 -0.2503 0.0169 0.2365 1.0997
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6581 0.1504 11.03 < 2e-16 ***
## FirstAuthorFemale1 0.2025 0.0787 2.57 0.01158 *
## LastAuthorFemale1 0.0596 0.0845 0.71 0.48232
## Year1998 -0.5444 0.4327 -1.26 0.21129
## Year1999 -0.6457 0.2416 -2.67 0.00877 **
## Year2000 -0.7103 0.1876 -3.79 0.00026 ***
## Year2001 -0.5001 0.1504 -3.33 0.00123 **
## Year2002 -0.3922 0.1723 -2.28 0.02490 *
## Year2003 -0.7363 0.2170 -3.39 0.00099 ***
## Year2004 -0.5604 0.2014 -2.78 0.00643 **
## Year2005 -0.7950 0.1631 -4.87 4.1e-06 ***
## Year2006 -0.6810 0.1689 -4.03 0.00011 ***
```

```

## Year2007          -0.7015      0.1980    -3.54  0.00060 ***
## Year2008          -0.6559      0.1967    -3.33  0.00120 **
## Year2009          -0.9157      0.2047    -4.47  2.0e-05 ***
## Year2010          -0.8359      0.2347    -3.56  0.00057 ***
## Year2011          -0.7716      0.1924    -4.01  0.00012 ***
## Year2012          -0.9763      0.2166    -4.51  1.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.385
## Multiple R-squared:  0.244, Adjusted R-squared:  0.116
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 105 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.394  0.895  0.942  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      8.40e-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 2.44 1      1.562
## Year              2.44 15      1.030

```

Residuals from first author



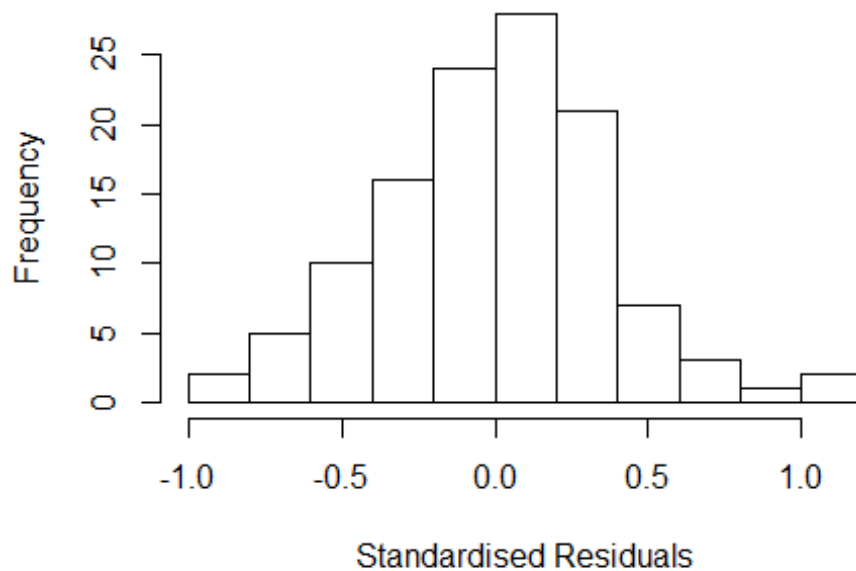
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9155 -0.2507 0.0249 0.2341 1.0685
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6739 0.1429 11.72 < 2e-16 ***
## FirstAuthorFemale1 0.2058 0.0775 2.66 0.00913 **
## Year1998 -0.5638 0.4282 -1.32 0.19091
## Year1999 -0.6460 0.2386 -2.71 0.00795 **
## Year2000 -0.7096 0.1835 -3.87 0.00019 ***
## Year2001 -0.5159 0.1429 -3.61 0.00047 ***
## Year2002 -0.3984 0.1679 -2.37 0.01950 *
## Year2003 -0.7411 0.2084 -3.56 0.00057 ***
## Year2004 -0.5772 0.1913 -3.02 0.00322 **
## Year2005 -0.7977 0.1582 -5.04 2.0e-06 ***
## Year2006 -0.6702 0.1680 -3.99 0.00013 ***
## Year2007 -0.6999 0.1928 -3.63 0.00044 ***
```

```

## Year2008          -0.6738      0.1892    -3.56  0.00056 ***
## Year2009          -0.9083      0.2013    -4.51  1.7e-05 ***
## Year2010          -0.8122      0.2424    -3.35  0.00113 **
## Year2011          -0.7727      0.1882    -4.11  8.1e-05 ***
## Year2012          -0.9642      0.2115    -4.56  1.4e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.387
## Multiple R-squared:  0.238, Adjusted R-squared:  0.119
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 103 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.426  0.889  0.943  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      8.40e-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 2.54 1          1.594
## Year            2.54 15          1.032

```

Residuals from last author



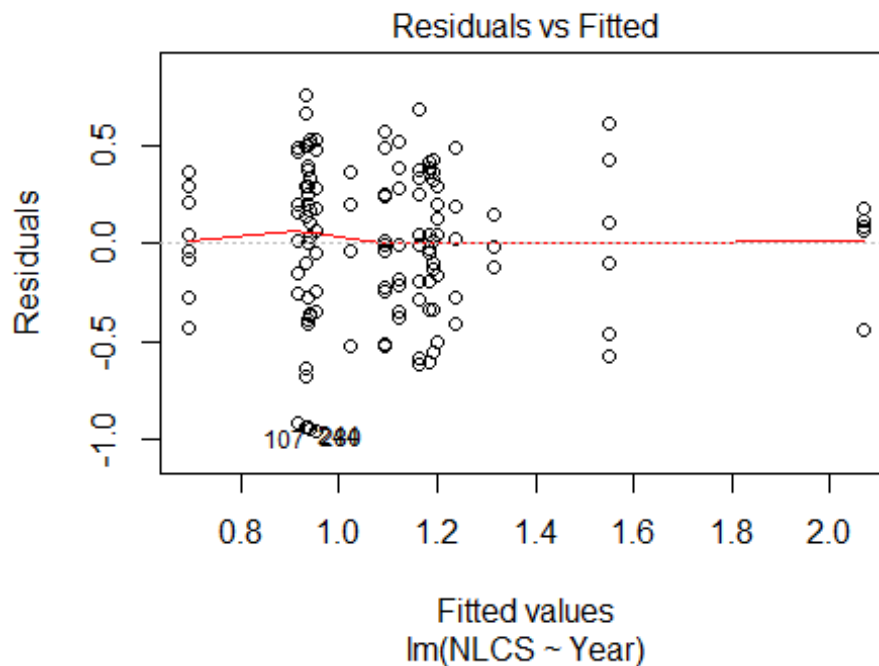
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.93613 -0.24751 -0.00244 0.29554 1.16712
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7076 0.1207 14.15 < 2e-16 ***
## LastAuthorFemale1 0.0670 0.0951 0.70 0.48313
## Year1998 -0.5255 0.4828 -1.09 0.27903
## Year1999 -0.6971 0.2196 -3.17 0.00198 **
## Year2000 -0.7358 0.1719 -4.28 4.2e-05 ***
## Year2001 -0.5496 0.1207 -4.55 1.5e-05 ***
## Year2002 -0.3753 0.1382 -2.72 0.00774 **
## Year2003 -0.6691 0.1984 -3.37 0.00105 **
## Year2004 -0.5141 0.1681 -3.06 0.00284 **
## Year2005 -0.7856 0.1478 -5.32 6.3e-07 ***
## Year2006 -0.6906 0.1541 -4.48 1.9e-05 ***
## Year2007 -0.6165 0.1683 -3.66 0.00040 ***
```

```

## Year2008          -0.5788      0.2054    -2.82   0.00581 **
## Year2009          -0.9327      0.1736    -5.37   4.9e-07 ***
## Year2010          -0.8492      0.2426    -3.50   0.00069 ***
## Year2011          -0.7332      0.1827    -4.01   0.00011 ***
## Year2012          -0.8907      0.1980    -4.50   1.8e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.417
## Multiple R-squared:  0.203, Adjusted R-squared:  0.0778
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.413  0.898  0.949  0.920  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      8.40e-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: 119"
## [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
##    9   13    9   13   11    4   11   15   17   15   14   12   16   16   16
## 2011 2012
##    9    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    4    8    6    3    9   10   10   10   11    8    9    9   11
## 2011 2012
##    5    9

```

```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    5    4    8    5    3    7   10    8   10    8    7    8    8   11
## 2011 2012
##    4    8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.6
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.5, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 4.44128606984584"
## [1] "Male first author team size 2018 geometric mean: 2.11474252688113"

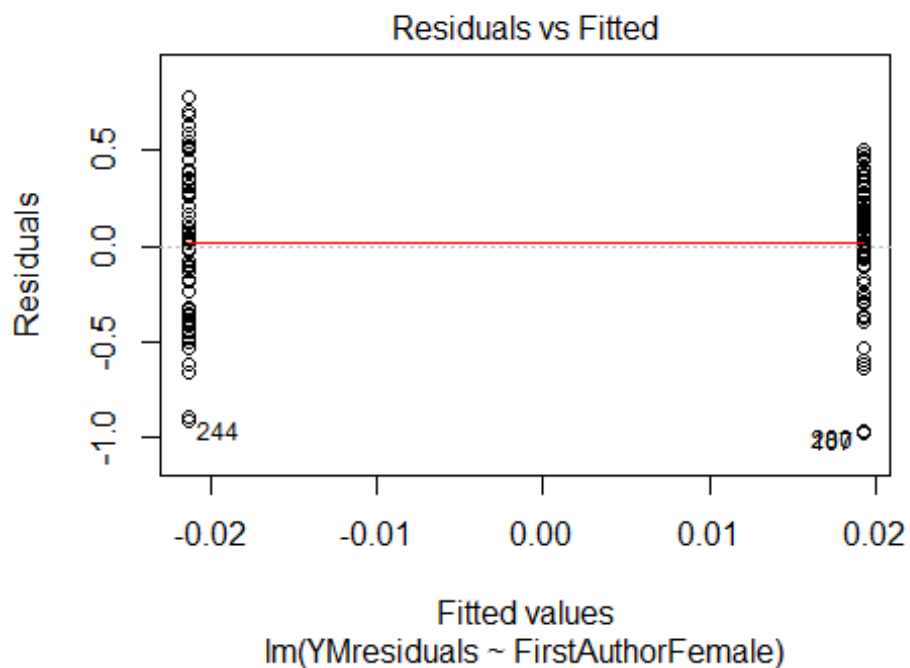
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
```



```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4"
## [1] "Male last author team size 2018 geometric mean: 3.1050186838948"

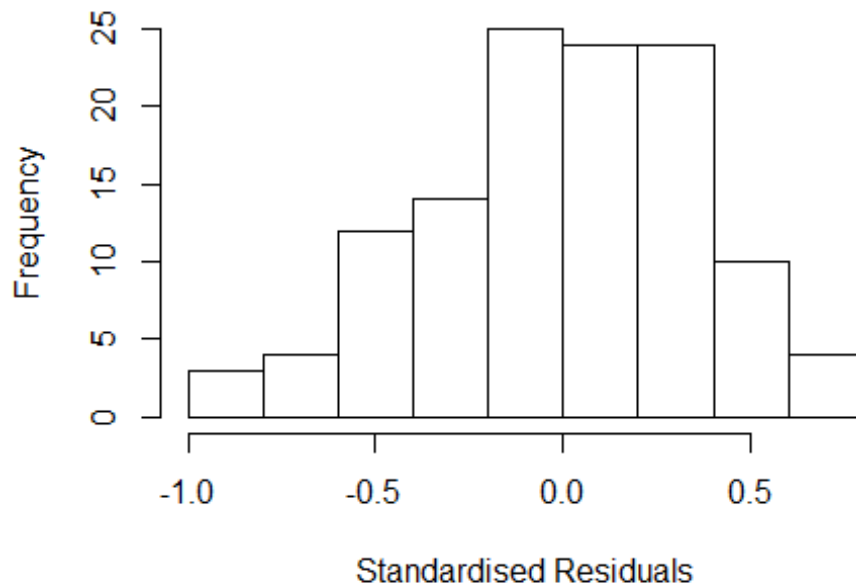
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, 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"
##
```

	GVI	F	Df	GVI ^{1/(2*Df)}
## FirstAuthorFemale	2.688	1		1.639
## LastAuthorFemale	1.768	1		1.330
## UniqueAuthors	10.316	4		1.339
## Year	22.100	16		1.102

Residuals from first and last author and team size



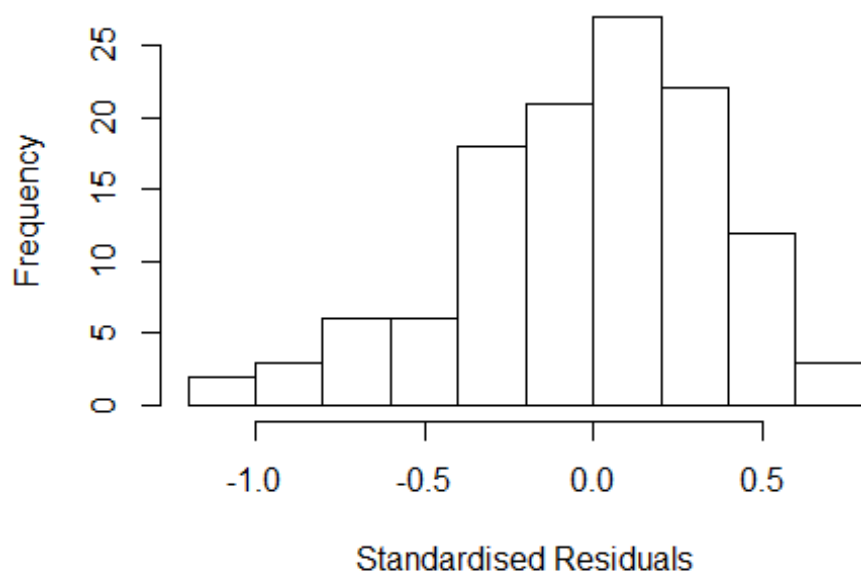
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9391 -0.2382 0.0103 0.2546 0.6810
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3946 0.3211 4.34 3.5e-05 ***
## FirstAuthorFemale1 -0.1926 0.0974 -1.98 0.05078 .
## LastAuthorFemale1 -0.0802 0.0942 -0.85 0.39658
## UniqueAuthors2 0.3844 0.1559 2.46 0.01546 *
## UniqueAuthors3 0.2363 0.1637 1.44 0.15209
## UniqueAuthors4 0.4322 0.1348 3.21 0.00182 **
## UniqueAuthors5 0.4647 0.1533 3.03 0.00312 **
## Year1997 0.3882 0.2839 1.37 0.17472
## Year1998 -0.7219 0.3096 -2.33 0.02178 *
## Year1999 -0.6437 0.2943 -2.19 0.03114 *
```

```

## Year2000          -0.5343      0.2903    -1.84    0.06873 .
## Year2001          -0.2429      0.3149    -0.77    0.44235
## Year2002          -0.3987      0.3017    -1.32    0.18947
## Year2003          -0.7452      0.3228    -2.31    0.02312 *
## Year2004          -0.3965      0.2977    -1.33    0.18603
## Year2005          -0.4213      0.3322    -1.27    0.20771
## Year2006          -0.4458      0.3163    -1.41    0.16193
## Year2007          -0.3802      0.3040    -1.25    0.21412
## Year2008          -0.5669      0.3489    -1.63    0.10740
## Year2009          -1.0502      0.2949    -3.56    0.00057 ***
## Year2010          -0.6473      0.2842    -2.28    0.02497 *
## Year2011          -0.3184      0.3104    -1.03    0.30760
## Year2012          -0.6431      0.3265    -1.97    0.05169 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.381
## Multiple R-squared:  0.426, Adjusted R-squared:  0.296
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 110 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.522  0.888  0.953  0.917  0.988  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           8.33e-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 1.871 1           1.368
## LastAuthorFemale  1.800 1           1.342
## Year              3.210 16           1.037

```

Residuals from first and last author



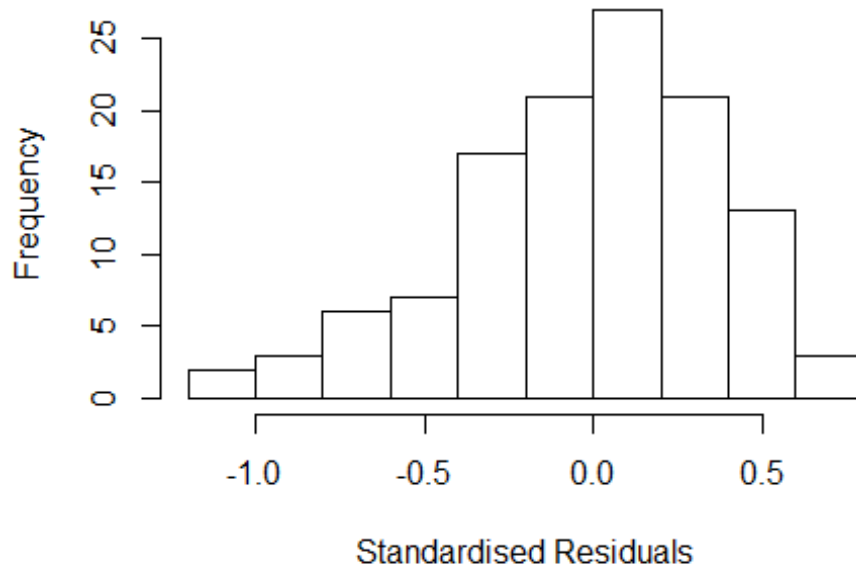
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.026 -0.307 0.029 0.287 0.734
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6572 0.2301 7.20 1.1e-10 ***
## FirstAuthorFemale1 -0.1109 0.0905 -1.23 0.2233
## LastAuthorFemale1 -0.0684 0.1123 -0.61 0.5439
## Year1997 0.4739 0.2319 2.04 0.0436 *
## Year1998 -0.5798 0.2730 -2.12 0.0361 *
## Year1999 -0.6178 0.2467 -2.50 0.0139 *
## Year2000 -0.4533 0.2423 -1.87 0.0643 .
## Year2001 -0.3061 0.2253 -1.36 0.1772
## Year2002 -0.3901 0.2371 -1.65 0.1030
## Year2003 -0.6576 0.2647 -2.48 0.0146 *
## Year2004 -0.3872 0.2410 -1.61 0.1112
## Year2005 -0.4340 0.2713 -1.60 0.1128
```

```

## Year2006          -0.4578      0.2616    -1.75    0.0831 .
## Year2007          -0.3238      0.2595    -1.25    0.2151
## Year2008          -0.5450      0.3800    -1.43    0.1545
## Year2009          -0.9798      0.2420    -4.05    0.0001 ***
## Year2010          -0.6133      0.2440    -2.51    0.0135 *
## Year2011          -0.2588      0.2642    -0.98    0.3295
## Year2012          -0.6317      0.2985    -2.12    0.0368 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.413
## Multiple R-squared:  0.339, Adjusted R-squared:  0.222
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 108 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.518  0.898  0.946  0.917  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      8.33e-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 1.873 1      1.368
## Year              1.873 16      1.020

```

Residuals from first author



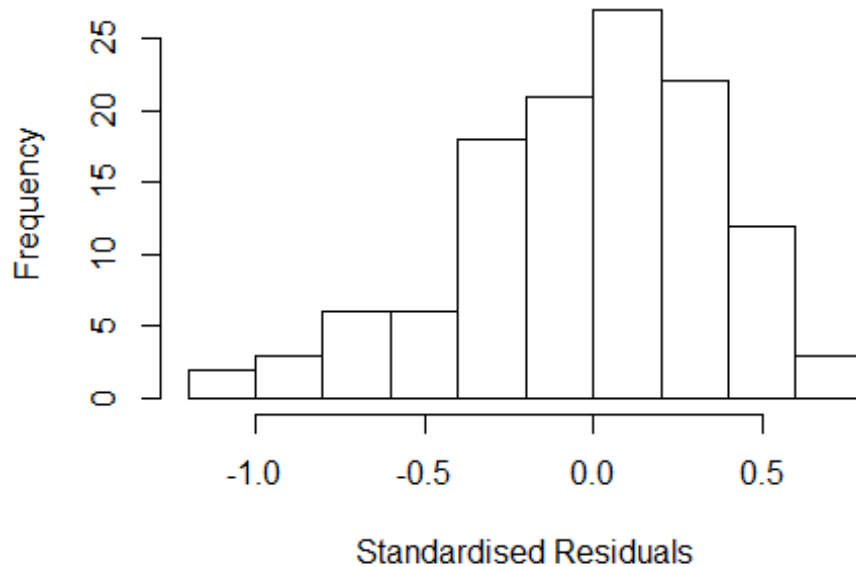
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0049 -0.3060 0.0279 0.2784 0.7369
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6354 0.2291 7.14 1.4e-10 ***
## FirstAuthorFemale1 -0.1177 0.0907 -1.30 0.19763
## Year1997 0.4841 0.2347 2.06 0.04172 *
## Year1998 -0.5749 0.2724 -2.11 0.03722 *
## Year1999 -0.6118 0.2536 -2.41 0.01763 *
## Year2000 -0.4427 0.2423 -1.83 0.07062 .
## Year2001 -0.2820 0.2229 -1.27 0.20866
## Year2002 -0.3870 0.2437 -1.59 0.11535
## Year2003 -0.6336 0.2644 -2.40 0.01837 *
## Year2004 -0.3618 0.2383 -1.52 0.13194
## Year2005 -0.4086 0.2705 -1.51 0.13399
## Year2006 -0.4312 0.2597 -1.66 0.09991 .
```

```

## Year2007          -0.3042      0.2561   -1.19  0.23768
## Year2008          -0.5179      0.4080   -1.27  0.20712
## Year2009          -0.9573      0.2405   -3.98  0.00013 ***
## Year2010          -0.6139      0.2521   -2.43  0.01664 *
## Year2011          -0.2314      0.2606   -0.89  0.37655
## Year2012          -0.6304      0.3017   -2.09  0.03914 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.404
## Multiple R-squared:  0.339, Adjusted R-squared:  0.228
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 107 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.517  0.899  0.941  0.912  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      8.33e-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 1.617 1      1.272
## Year            1.617 16      1.015

```

Residuals from last author



```
## [1] "List of 0 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
## -1.0215 -0.2851 0.0329 0.2870 0.6785
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5833 0.2099 7.54 2e-11 ***
## LastAuthorFemale1 -0.0862 0.1154 -0.75 0.45661
## Year1997 0.5101 0.2268 2.25 0.02664 *
## Year1998 -0.5224 0.2733 -1.91 0.05873 .
## Year1999 -0.6033 0.2435 -2.48 0.01486 *
## Year2000 -0.4170 0.2346 -1.78 0.07846 .
## Year2001 -0.2691 0.2195 -1.23 0.22302
## Year2002 -0.3659 0.2341 -1.56 0.12118
## Year2003 -0.5950 0.2438 -2.44 0.01638 *
## Year2004 -0.3694 0.2386 -1.55 0.12472
## Year2005 -0.4159 0.2521 -1.65 0.10217
## Year2006 -0.4649 0.2496 -1.86 0.06541 .
```

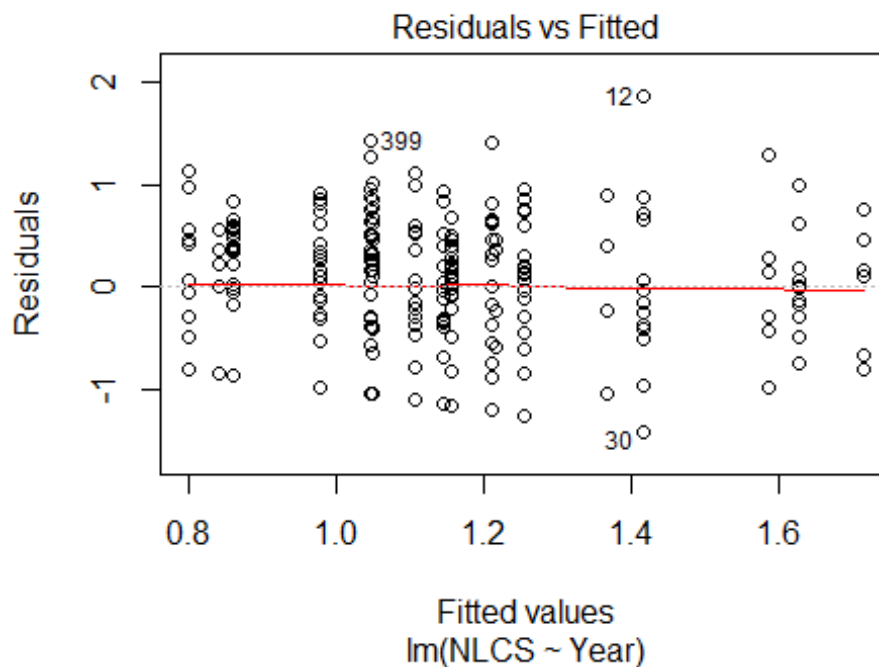


```

## Year2007          -0.3590      0.2554   -1.41  0.16287
## Year2008          -0.5618      0.3769   -1.49  0.13911
## Year2009          -0.9181      0.2282   -4.02  0.00011 ***
## Year2010          -0.5787      0.2381   -2.43  0.01681 *
## Year2011          -0.2706      0.2715   -1.00  0.32123
## Year2012          -0.5706      0.2773   -2.06  0.04216 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.429
## Multiple R-squared:  0.325, Adjusted R-squared:  0.212
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 106 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.550  0.905   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      8.33e-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: 120"
## [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
##    5   25   18   17   21   21    9   19   16   20   23   28   33   24   25
## 2011 2012
##   36   33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4   13    6    4    6   11    7   16   14   16   17   24   29   22   22
## 2011 2012

```

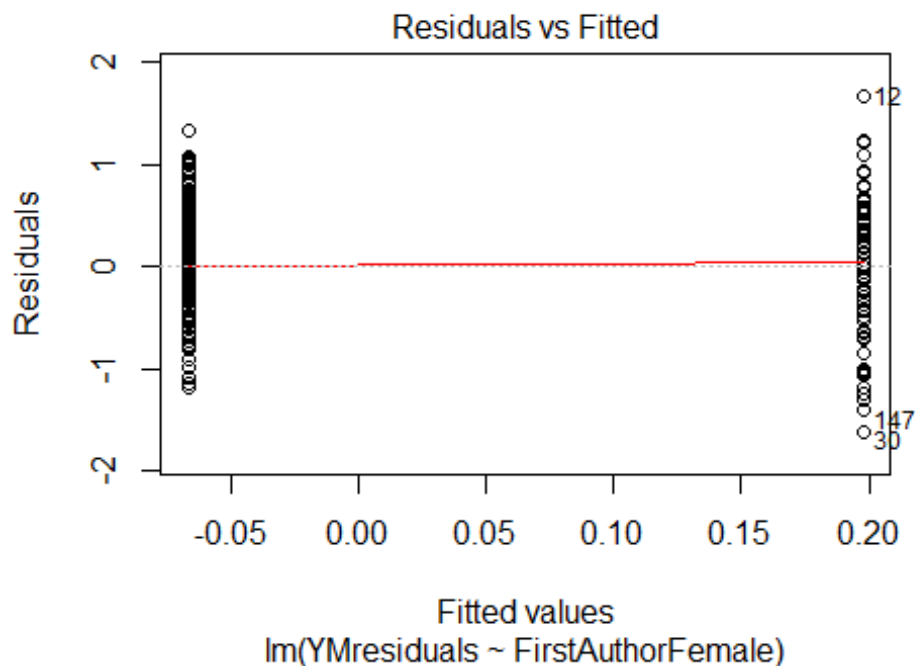
```
## 30 28
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 12 6 4 6 9 7 16 13 14 16 22 27 19 20
## 2011 2012
## 27 25
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 16, df = 16, p-value = 0.4
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.9, df = 1, p-value = 0.05
## [1] "Female first author team size 2018 geometric mean: 2.27494983790605"
## [1] "Male first author team size 2018 geometric mean: 4.37344829577311"
## 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 = 38, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.40657979627694"
## [1] "Male last author team size 2018 geometric mean: 4"

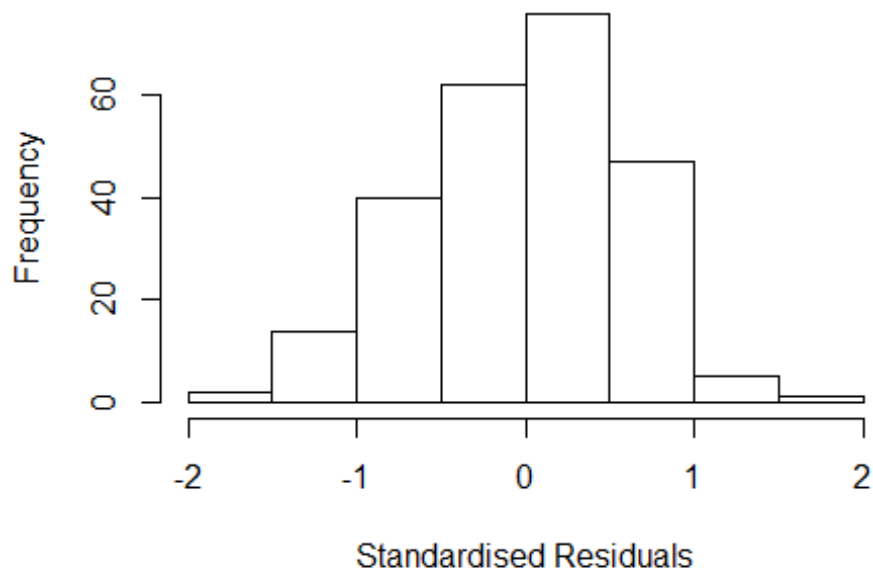
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16, 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	1.337	1	1.156
LastAuthorFemale	1.742	1	1.320
UniqueAuthors	2.869	4	1.141
Year	3.880	16	1.043

Residuals from first and last author and team size



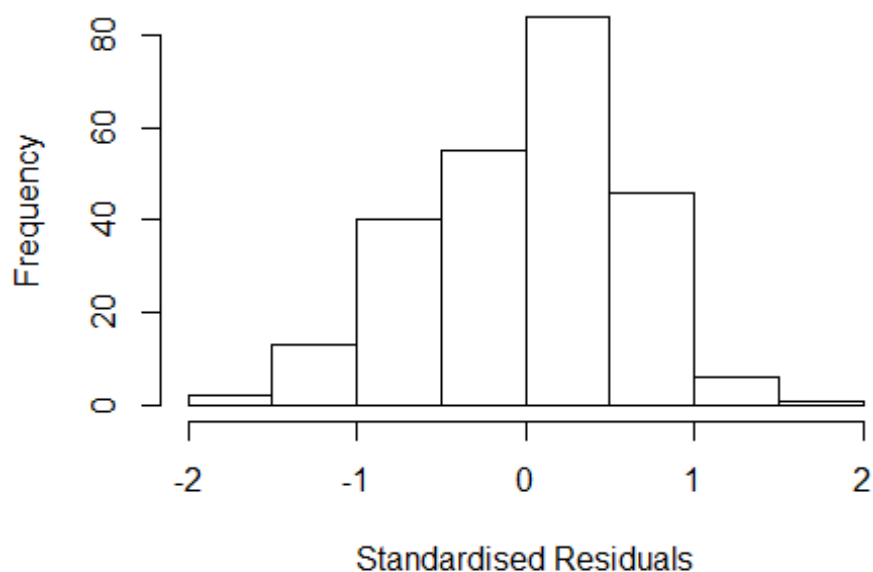
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.6397 -0.4086 0.0463 0.4056 1.5321
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.59969 0.26747 5.98 8.7e-09 ***
## FirstAuthorFemale1 -0.26682 0.10549 -2.53 0.0121 *
## LastAuthorFemale1 -0.11495 0.10526 -1.09 0.2760
## UniqueAuthors2 0.00943 0.10705 0.09 0.9299
## UniqueAuthors3 0.24608 0.11272 2.18 0.0301 *
## UniqueAuthors4 0.53201 0.19439 2.74 0.0067 **
## UniqueAuthors5 0.48334 0.21445 2.25 0.0252 *
## Year1997 -0.10502 0.30345 -0.35 0.7296
## Year1998 0.10003 0.33774 0.30 0.7674
## Year1999 -0.17533 0.56139 -0.31 0.7551
```

```

## Year2000      0.34503      0.34118      1.01      0.3130
## Year2001      0.30194      0.31202      0.97      0.3342
## Year2002     -0.42742      0.32237     -1.33      0.1862
## Year2003     -0.11020      0.32433     -0.34      0.7344
## Year2004     -0.51924      0.31240     -1.66      0.0979 .
## Year2005     -0.22414      0.28574     -0.78      0.4336
## Year2006     -0.35573      0.27688     -1.28      0.2002
## Year2007     -0.19848      0.26011     -0.76      0.4462
## Year2008     -0.59019      0.26832     -2.20      0.0289 *
## Year2009     -0.42686      0.27330     -1.56      0.1197
## Year2010     -0.05030      0.26325     -0.19      0.8486
## Year2011     -0.49416      0.26667     -1.85      0.0652 .
## Year2012     -0.44197      0.30092     -1.47      0.1433
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.609
## Multiple R-squared:  0.243, Adjusted R-squared:  0.169
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 229 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.448  0.880  0.949  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      4.05e-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 1.450 1      1.204
## LastAuthorFemale  1.488 1      1.220
## Year              1.562 16      1.014

```

Residuals from first and last author



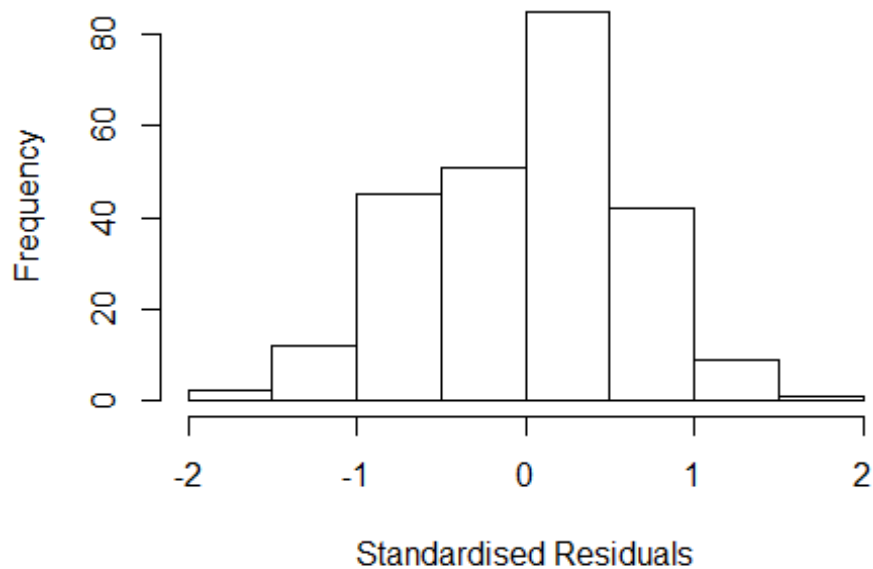
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5815 -0.4568 0.0713 0.4494 1.6915
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.6550 0.2663 6.21 2.4e-09 ***
## FirstAuthorFemale1 -0.2311 0.1192 -1.94 0.054 .
## LastAuthorFemale1 -0.2039 0.1063 -1.92 0.056 .
## Year1997 -0.0735 0.3200 -0.23 0.818
## Year1998 0.1716 0.3740 0.46 0.647
## Year1999 -0.1978 0.6392 -0.31 0.757
## Year2000 0.3145 0.3342 0.94 0.348
## Year2001 0.3207 0.2957 1.08 0.279
## Year2002 -0.4537 0.3152 -1.44 0.151
## Year2003 -0.0951 0.3183 -0.30 0.765
## Year2004 -0.5461 0.3077 -1.77 0.077 .
## Year2005 -0.2162 0.2804 -0.77 0.442
```

```

## Year2006          -0.1926      0.2675   -0.72    0.472
## Year2007          -0.1824      0.2590   -0.70    0.482
## Year2008          -0.4992      0.2624   -1.90    0.058 .
## Year2009          -0.3652      0.2715   -1.35    0.180
## Year2010           0.0562      0.2631    0.21    0.831
## Year2011          -0.4071      0.2624   -1.55    0.122
## Year2012          -0.2867      0.2953   -0.97    0.333
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.644
## Multiple R-squared:  0.185, Adjusted R-squared:  0.12
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 232 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.470  0.889  0.952  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      4.05e-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 1.248 1      1.117
## Year              1.248 16      1.007

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.5778 -0.4688 0.0532 0.4318 1.6952
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.578250 0.258472 6.11 4.3e-09 ***
## FirstAuthorFemale1 -0.358324 0.109271 -3.28 0.0012 **
## Year1997 -0.000492 0.318881 0.00 0.9988
## Year1998 0.234141 0.379636 0.62 0.5380
## Year1999 -0.126904 0.606296 -0.21 0.8344
## Year2000 0.348122 0.334223 1.04 0.2987
## Year2001 0.346849 0.291975 1.19 0.2361
## Year2002 -0.406879 0.316909 -1.28 0.2005
## Year2003 -0.058419 0.317166 -0.18 0.8540
## Year2004 -0.506774 0.303493 -1.67 0.0963 .
## Year2005 -0.208699 0.276509 -0.75 0.4512
## Year2006 -0.198446 0.268710 -0.74 0.4610
```



```

## Year2007          -0.117067    0.253737   -0.46    0.6450
## Year2008          -0.474456    0.264529   -1.79    0.0742 .
## Year2009          -0.307276    0.273138   -1.12    0.2618
## Year2010           0.094385    0.264764    0.36    0.7218
## Year2011          -0.398170    0.262408   -1.52    0.1306
## Year2012          -0.274988    0.298103   -0.92    0.3573
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.647
## Multiple R-squared:  0.173, Adjusted R-squared:  0.111
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 224 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.472  0.883  0.947  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      4.05e-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 1.283  1      1.133
## 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.4774 -0.4583  0.0542  0.4539  1.7956

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5492    0.2540   6.10 4.5e-09 ***
## LastAuthorFemale1 -0.3293    0.0993  -3.32  0.0011 **
## Year1997          -0.0718    0.3300  -0.22  0.8280
## Year1998           0.1693    0.3715   0.46  0.6491
## Year1999          -0.1008    0.5923  -0.17  0.8650
## Year2000           0.3606    0.3344   1.08  0.2820
## Year2001           0.3569    0.2904   1.23  0.2204
## Year2002          -0.4580    0.3171  -1.44  0.1501
## Year2003          -0.0917    0.3186  -0.29  0.7737
## Year2004          -0.5146    0.3116  -1.65  0.1000 .
## Year2005          -0.1952    0.2793  -0.70  0.4853
## Year2006          -0.1421    0.2708  -0.52  0.6002
## Year2007          -0.1887    0.2583  -0.73  0.4658
## Year2008          -0.4781    0.2630  -1.82  0.0704 .
## Year2009          -0.3635    0.2740  -1.33  0.1860
## Year2010           0.0486    0.2611   0.19  0.8526
## Year2011          -0.3765    0.2623  -1.44  0.1525
## Year2012          -0.2866    0.2963  -0.97  0.3344
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.65
## Multiple R-squared:  0.169, Adjusted R-squared:  0.107
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 222 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.426  0.883   0.950   0.916   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      4.05e-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: 247"
## [1] ""

```

```

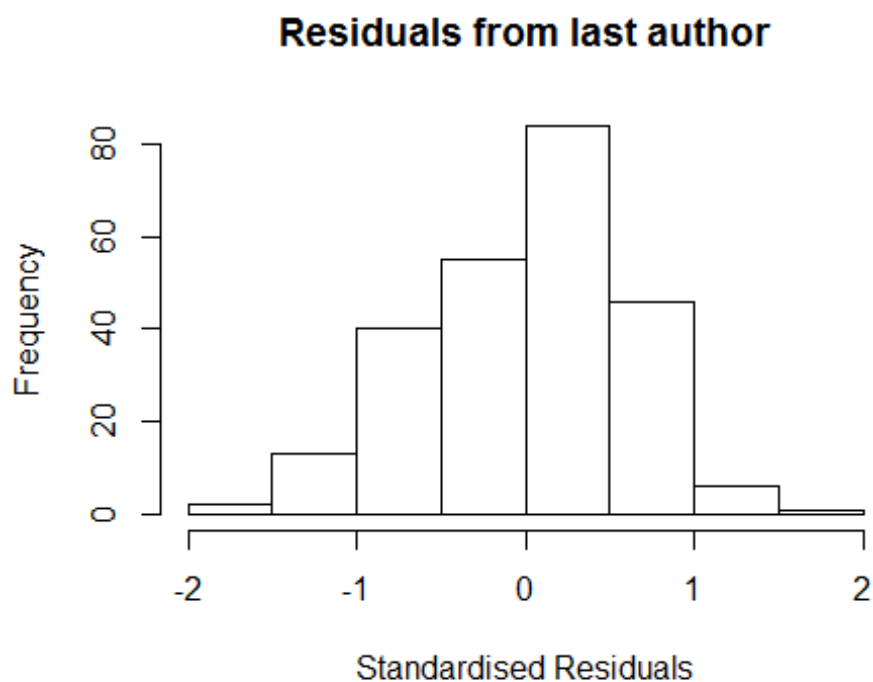
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2901"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2004 2007 2008 2009 2010 2011 2012
##    3    1    2    4    1    4    3
##
## 2004 2007 2008 2009 2010 2011 2012
##    1    1    1    1    0    1    3
##
## 2004 2007 2008 2009 2010 2011 2012
##    1    1    1    1    0    1    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5"
## [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: 7"
## [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]"
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    3    2    4    4    5    1    6    2    4    9    8    6    7    9    6
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    1    1    2    5    1    6    1    3    5    6    5    6    7    5
##
## 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    1    1    2    4    0    6    1    3    5    5    2    6    4    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male first author team size 2018 geometric mean: 3.87298334620742"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

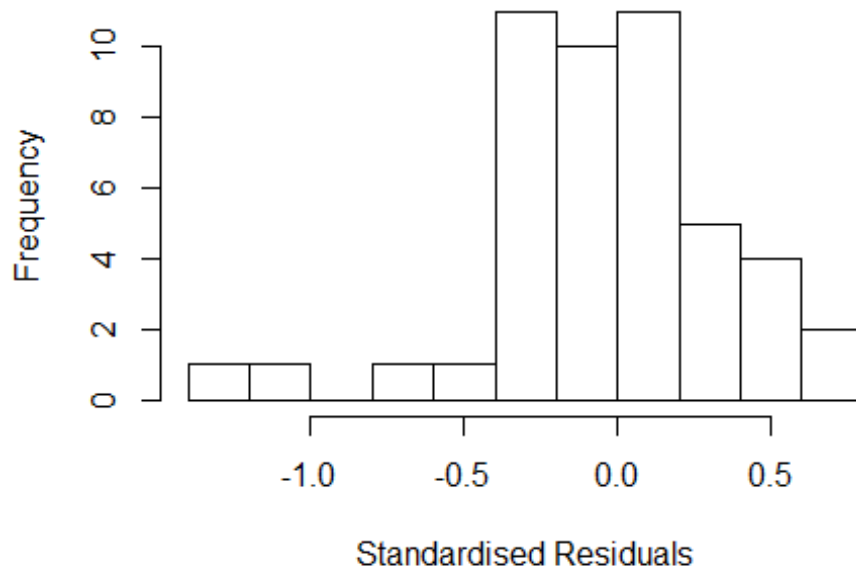
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.44948974278318"
## [1] "Male last author team size 2018 geometric mean: 3.87298334620742"

## 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 = 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 1.880e+14 1      1.371e+07
## LastAuthorFemale -1.816e+01 1           NaN
## UniqueAuthors    4.527e+31 4      9.057e+03
## Year              3.745e+32 13     1.790e+01
```

Residuals from first and last author and team size



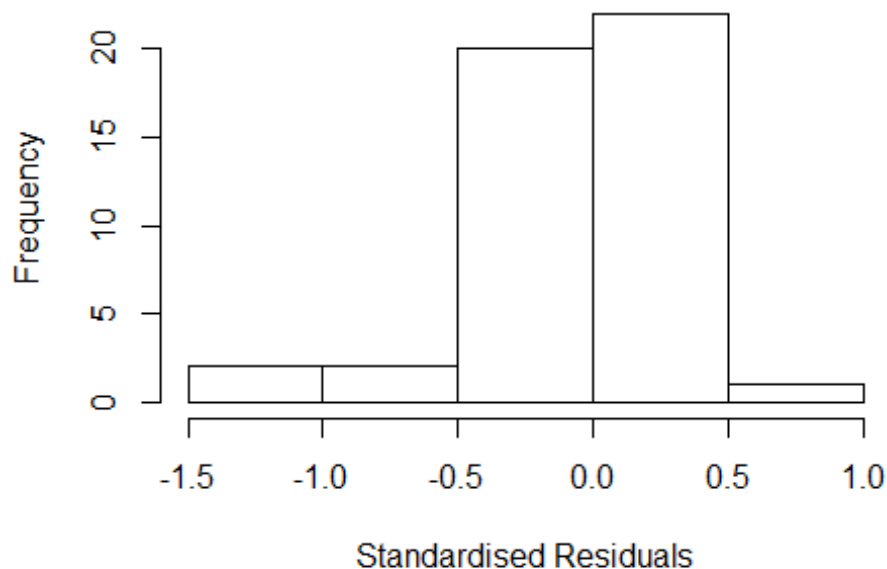
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22e+00 -2.21e-01 7.77e-16 1.74e-01 6.96e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.0496 0.4837 4.24 0.00024 ***
## FirstAuthorFemale1 -0.0419 0.1153 -0.36 0.71894
## LastAuthorFemale1 0.0403 0.1789 0.23 0.82330
## UniqueAuthors2 0.1003 0.2342 0.43 0.67169
## UniqueAuthors3 -0.3140 0.4032 -0.78 0.44288
## UniqueAuthors4 0.6090 0.3349 1.82 0.08016 .
## UniqueAuthors5 0.2173 0.3188 0.68 0.50135
## Year1998 -0.7078 0.5363 -1.32 0.19796
## Year2000 0.2503 0.3294 0.76 0.45382
## Year2001 -0.1500 0.3189 -0.47 0.64198
```

```

## Year2002          -0.9727      0.4850    -2.01  0.05504 .
## Year2004          -0.8308      0.6155    -1.35  0.18825
## Year2005          -1.0966      0.4837    -2.27  0.03160 *
## Year2006          -1.1514      0.4808    -2.39  0.02385 *
## Year2007          -1.1013      0.4798    -2.30  0.02970 *
## Year2008          -1.0930      0.5544    -1.97  0.05900 .
## Year2009          -1.5329      0.5533    -2.77  0.01000 *
## Year2010          -1.0782      0.4231    -2.55  0.01683 *
## Year2011          -0.8989      0.5482    -1.64  0.11267
## Year2012          -1.3669      0.4604    -2.97  0.00620 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.455
## Multiple R-squared:  0.482, Adjusted R-squared:  0.117
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.454  0.931  0.966  0.930  0.990  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.13e-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 -7.450e+14  1      NaN
## LastAuthorFemale  5.983e+00  1      2.446
## Year              7.524e+03 13      1.410

```

Residuals from first and last author



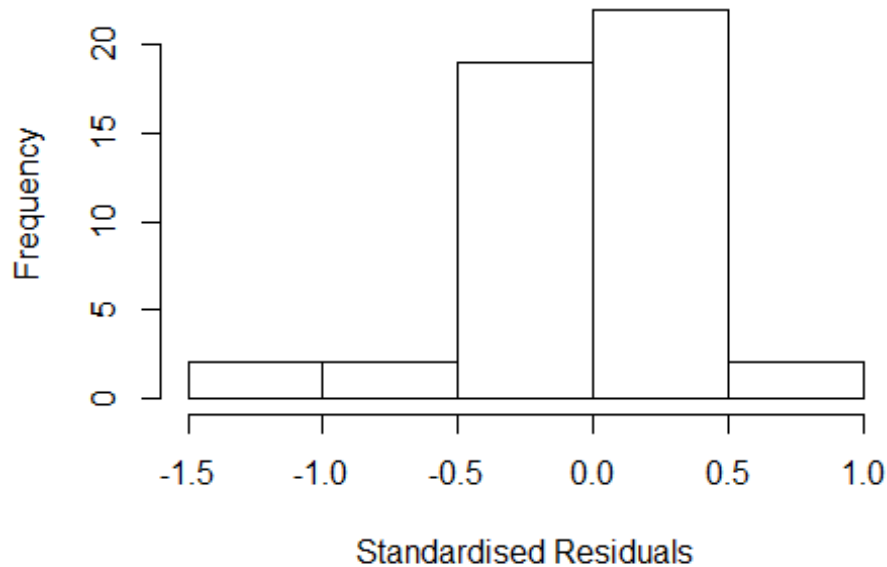
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.48e+00 -2.58e-01 -5.55e-16 2.76e-01 5.91e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.81e+00 3.36e-01 5.38 7.2e-06 ***
## FirstAuthorFemale1 8.94e-05 1.74e-01 0.00 0.9996
## LastAuthorFemale1 -7.33e-02 1.78e-01 -0.41 0.6835
## Year1998 -2.48e-01 3.36e-01 -0.74 0.4654
## Year2000 1.37e-01 3.33e-01 0.41 0.6839
## Year2001 -1.29e-01 3.34e-01 -0.39 0.7017
## Year2002 -6.01e-01 2.98e-01 -2.01 0.0527 .
## Year2004 -2.59e-01 4.07e-01 -0.64 0.5297
## Year2005 -8.54e-01 3.36e-01 -2.54 0.0162 *
## Year2006 -7.52e-01 3.21e-01 -2.34 0.0258 *
## Year2007 -6.55e-01 3.01e-01 -2.18 0.0374 *
## Year2008 -6.89e-01 4.48e-01 -1.54 0.1339
```

```

## Year2009          -1.15e+00  3.95e-01  -2.92  0.0065 **
## Year2010          -9.00e-01  3.54e-01  -2.54  0.0163 *
## Year2011          -4.94e-01  3.15e-01  -1.57  0.1274
## Year2012          -1.06e+00  3.61e-01  -2.93  0.0063 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.421, Adjusted R-squared:  0.141
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.280  0.935  0.962  0.923  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      2.13e-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 -1.072e+15  1          NaN
## Year              -1.072e+15 13          NaN

```


Residuals from first author



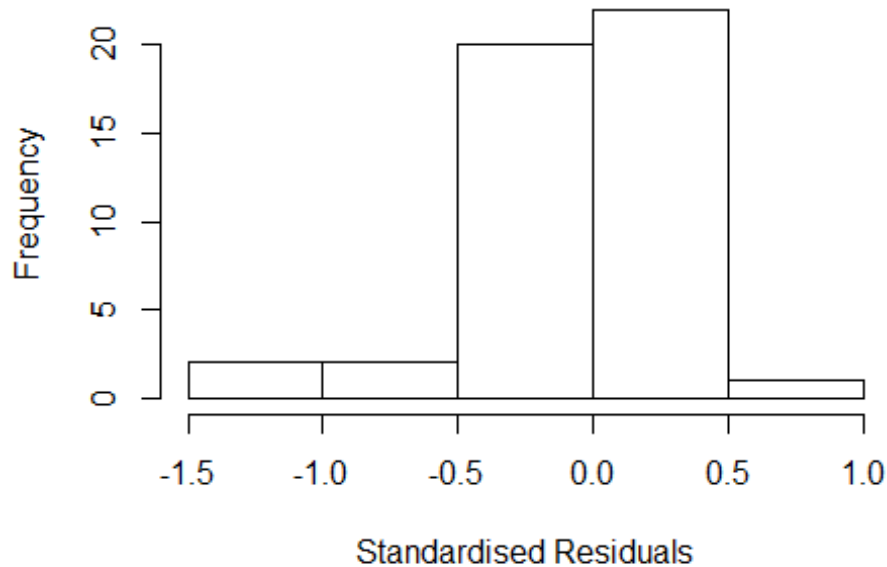
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.48107 -0.25628 0.00395 0.26239 0.65849
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.768 0.298 5.92 1.4e-06 ***
## FirstAuthorFemale1 -0.034 0.141 -0.24 0.8111
## Year1998 -0.209 0.298 -0.70 0.4889
## Year2000 0.210 0.270 0.78 0.4429
## Year2001 -0.146 0.311 -0.47 0.6420
## Year2002 -0.590 0.288 -2.05 0.0487 *
## Year2004 -0.253 0.363 -0.70 0.4913
## Year2005 -0.815 0.298 -2.73 0.0102 *
## Year2006 -0.740 0.305 -2.42 0.0212 *
## Year2007 -0.640 0.288 -2.22 0.0338 *
## Year2008 -0.691 0.401 -1.72 0.0946 .
## Year2009 -1.097 0.341 -3.22 0.0030 **
```

```

## Year2010          -0.893      0.335   -2.67   0.0118 *
## Year2011          -0.468      0.315   -1.49   0.1469
## Year2012          -1.052      0.364   -2.89   0.0068 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.485
## Multiple R-squared:  0.409, Adjusted R-squared:  0.151
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.331  0.938  0.963  0.930  0.985  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      2.13e-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 -3.694 1      NaN
## Year             -3.694 13      NaN

```

Residuals from last author



```
## [1] "List of 0 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 -2.57e-01 3.33e-16 2.77e-01 5.86e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.8137 0.3175 5.71 2.5e-06 ***
## LastAuthorFemale1 -0.0797 0.1543 -0.52 0.6089
## Year1998 -0.2547 0.3175 -0.80 0.4284
## Year2000 0.1303 0.3175 0.41 0.6843
## Year2001 -0.1290 0.3132 -0.41 0.6831
## Year2002 -0.6021 0.2917 -2.06 0.0472 *
## Year2004 -0.2789 0.3811 -0.73 0.4696
## Year2005 -0.8607 0.3175 -2.71 0.0107 *
## Year2006 -0.7543 0.3177 -2.37 0.0237 *
## Year2007 -0.6578 0.2942 -2.24 0.0325 *
## Year2008 -0.6999 0.4236 -1.65 0.1083
## Year2009 -1.1592 0.3836 -3.02 0.0049 **
```

```

## Year2010          -0.9058      0.3402    -2.66    0.0120 *
## Year2011          -0.4959      0.3105    -1.60    0.1201
## Year2012          -1.0592      0.3575    -2.96    0.0057 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.48
## Multiple R-squared:  0.411, Adjusted R-squared:  0.154
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.338  0.938  0.965  0.927  0.983  0.997
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          2.13e-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: 47"
## [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]"
##
## 2001
##    1
##
## 2001
##    0
##
## 2001
##    0
## [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 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: 0"
## [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]"
##
## 2001 2008
## 1 1
##
## 2001 2008
## 0 1
##
## 2001 2008
## 0 1
## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 1"
## [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]"
##
## 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012
## 1 4 2 3 1 4 6 2 4 3 3 11
##
## 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012
## 1 3 2 2 1 4 4 2 4 3 2 9
##
## 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2011 2012
## 1 3 1 2 1 4 4 2 4 3 2 7
## [1] "Heteroscedasticity checks, confirming that there are problems with

```

```

these"
## [1] "Female first author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male first author team size 2018 geometric mean: 6"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4"
## [1] "Male last author team size 2018 geometric mean: 4.24264068711928"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, 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"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.973e+15  1      4.442e+07
## LastAuthorFemale  2.721e+14  1      1.650e+07
## Year              3.057e+29 11      2.189e+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
## -1.14e+00 -2.41e-01  7.49e-16  2.92e-01  1.63e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.03e+00   2.40e-08  4.30e+07  <2e-16 ***
## FirstAuthorFemale1  5.89e-02   3.44e-01  1.70e-01  0.8657
## LastAuthorFemale1  4.78e-01   3.52e-01  1.36e+00  0.1897
## Year2001         -1.29e+00   4.41e-01 -2.93e+00  0.0082 **
## Year2002         -1.16e+00   3.52e-01 -3.29e+00  0.0037 **
## Year2003         -1.11e+00   5.09e-01 -2.19e+00  0.0406 *
## Year2004          2.51e-01   3.44e-01  7.30e-01  0.4740

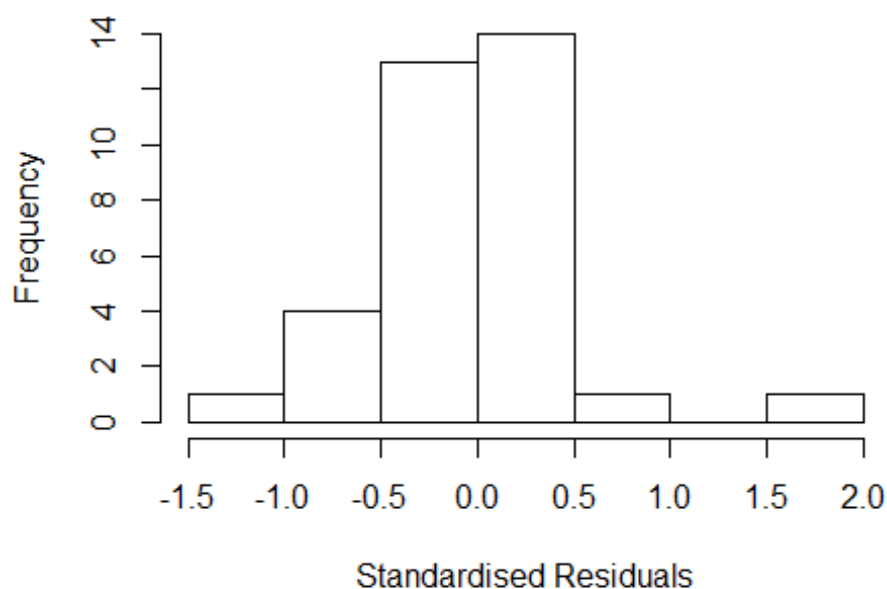
```

```

## Year2005          2.15e-01  2.33e-01  9.20e-01  0.3688
## Year2006          -8.78e-02  2.85e-01 -3.10e-01  0.7614
## Year2007          -3.40e-01  6.09e-01 -5.60e-01  0.5831
## Year2008          -5.98e-01  3.94e-01 -1.52e+00  0.1444
## Year2009          -1.41e-01  5.21e-01 -2.70e-01  0.7897
## Year2011           5.59e-01  4.71e-01  1.19e+00  0.2489
## Year2012          -4.24e-01  4.66e-01 -9.10e-01  0.3732
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.625
## Multiple R-squared:  0.469, Adjusted R-squared:  0.124
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.478  0.946  0.957  0.929  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      2.94e-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"
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

```

Residuals from first and last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 4.374e+14  1      2.091e+07
## Year              4.374e+14 11      4.629e+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 + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.0857 -0.4547  0.0375  0.2955  1.2083
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.03e+00   0.00e+00      Inf <2e-16 ***
## FirstAuthorFemale1  9.85e-02   3.41e-01   2.90e-01  0.7758
## Year2001        -1.02e+00   3.54e-01  -2.87e+00  0.0091 **
## Year2002        -6.80e-01   2.19e-08  -3.11e+07 <2e-16 ***
## Year2003        -6.76e-01   3.49e-01  -1.93e+00  0.0667 .
## Year2004         2.11e-01   3.41e-01   6.20e-01  0.5424
## Year2005         3.16e-01   2.92e-01   1.08e+00  0.2906
## Year2006         2.64e-01   1.64e-01   1.61e+00  0.1215
```



```

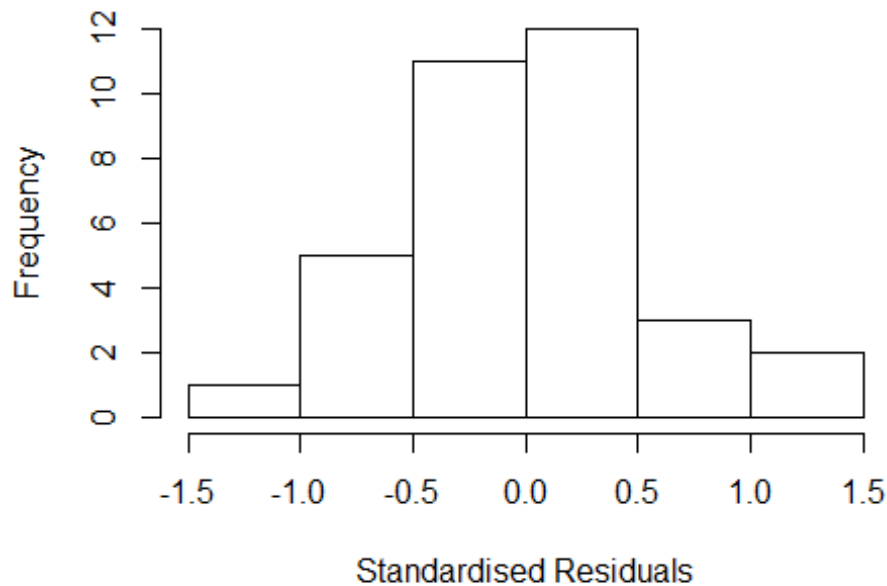
## Year2007          9.85e-02  4.90e-01  2.00e-01  0.8426
## Year2008         -4.29e-01  4.12e-01 -1.04e+00  0.3096
## Year2009          1.60e-01  5.46e-01  2.90e-01  0.7723
## Year2011          5.19e-01  4.76e-01  1.09e+00  0.2876
## Year2012         -4.48e-02  2.91e-01 -1.50e-01  0.8792
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.745
## Multiple R-squared:  0.381, Adjusted R-squared:  0.0267
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 29 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.775  0.941  0.966  0.949  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      2.94e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale NaN 1          NaN
## Year            NaN 11          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
## -1.097 -0.240  0.020  0.268  1.711
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.03e+00   1.92e-08  5.36e+07 < 2e-16 ***
## LastAuthorFemale1 5.14e-01   3.68e-01  1.40e+00  0.17690
## Year2001      -1.26e+00   2.79e-01 -4.52e+00  0.00019 ***
## Year2002      -1.19e+00   3.68e-01 -3.25e+00  0.00385 **
## Year2003      -1.09e+00   3.75e-01 -2.91e+00  0.00837 **
## Year2004       3.10e-01   0.00e+00      Inf < 2e-16 ***
## Year2005       2.22e-01   1.96e-01  1.13e+00  0.26938
## Year2006      -1.00e-01   2.83e-01 -3.50e-01  0.72620
```

```

## Year2007          -3.17e-01   4.83e-01 -6.60e-01   0.51940
## Year2008          -5.90e-01   4.51e-01 -1.31e+00   0.20445
## Year2009          -1.05e-01   3.81e-01 -2.80e-01   0.78446
## Year2011           6.18e-01   3.49e-01  1.77e+00   0.09077 .
## Year2012          -4.49e-01   4.66e-01 -9.60e-01   0.34584
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.574
## Multiple R-squared:  0.491, Adjusted R-squared:  0.2
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 29 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.354  0.942  0.958  0.922  0.995  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      2.94e-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: 34"
## [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]"
##
## 1999 2002 2003 2004 2005 2007 2008 2009 2010 2012
##    1    1    2    1    2    2    3    2    1    1
##
## 1999 2002 2003 2004 2005 2007 2008 2009 2010 2012
##    1    1    2    1    2    1    3    2    1    1
##
## 1999 2002 2003 2004 2005 2007 2008 2009 2010 2012
##    1    1    2    1    2    1    3    2    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with

```

```

these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 3"

## 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: 3"
## [1] "Male last author team size 2018 geometric mean: 3"

## 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"
## [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: 15"
## [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]"
##
## 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 1 1 1 2 1 1 1 1 1 3 2 1 3
##
## 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 0 1 0 1 0 1 1 1 0 3 2 1 3
##
## 1998 1999 2000 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 0 1 0 1 0 1 1 1 0 3 2 1 3
## [1] "Heteroscedasticity checks, confirming that there are problems with

```

```

these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [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

## [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): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

## [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: 14"
## [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]"
##
## 2002 2004 2006 2007 2008 2009 2010 2011
##    1    1    1    1    1    2    1    1
##
## 2002 2004 2006 2007 2008 2009 2010 2011
##    1    0    0    1    0    0    0    0
##
## 2002 2004 2006 2007 2008 2009 2010 2011
##    1    0    0    1    0    0    0    0
## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 2"
## [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]"
##
## 1998 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    1    3    1    3    1    2    2
##
## 1998 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    1    1    1    3    1    2    2
##
## 1998 2003 2005 2006 2007 2008 2009 2010 2011 2012
##    1    0    1    1    1    1    3    1    2    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.58341668418149"
## [1] "Male first author team size 2018 geometric mean: 3"

## 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 = 2.5, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.63214802590498"
## [1] "Male last author team size 2018 geometric mean: 2.63214802590498"

## 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] "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: 13"
## [1] ""
## [1] ""
## [1] "#####"

```

```

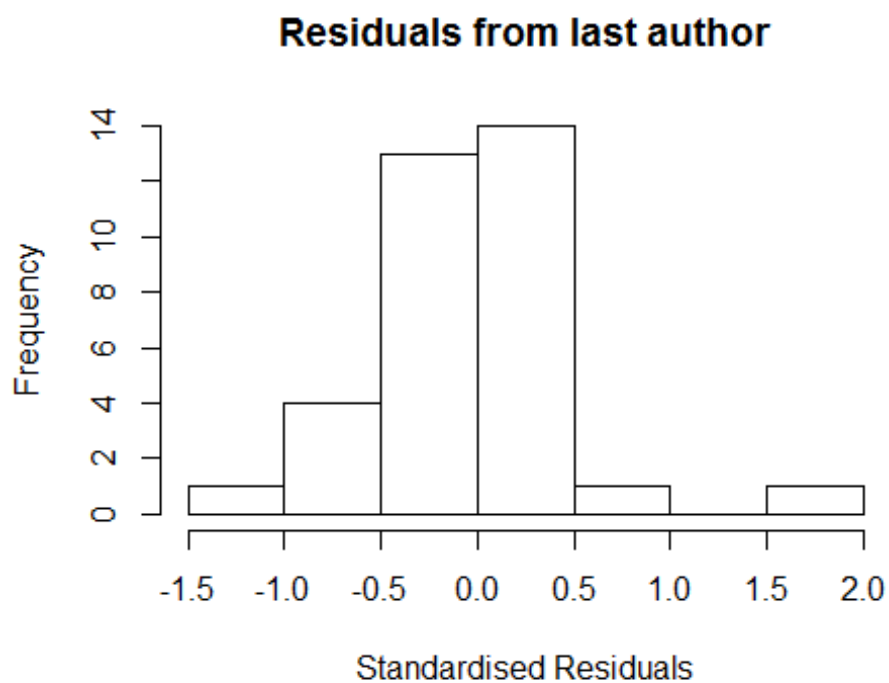
## [1] "Analysis of AJSC 2910"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    7    3    1    3    1    4    3    4    5    5    6    7    9    15
##
## 1997 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    5    1    1    2    0    3    2    4    3    1    4    7    8    12
##
## 1997 1999 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    5    1    1    2    0    3    1    4    3    1    4    7    8    12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.51984209978975"
## [1] "Male first author team size 2018 geometric mean: 1.41421356237309"

## 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 = 13, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.63901582154579"
## [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

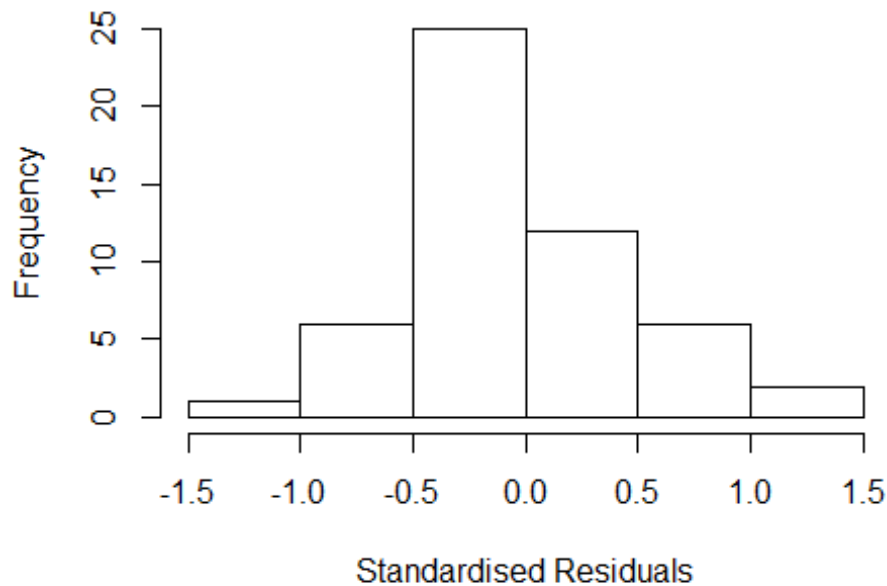
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 18, 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	3.934e+15	1	6.273e+07
## LastAuthorFemale	-6.183e+14	1	NaN
## UniqueAuthors	-1.482e+29	4	NaN
## Year	5.312e+43	12	6.636e+01

Residuals from first and last author and team size



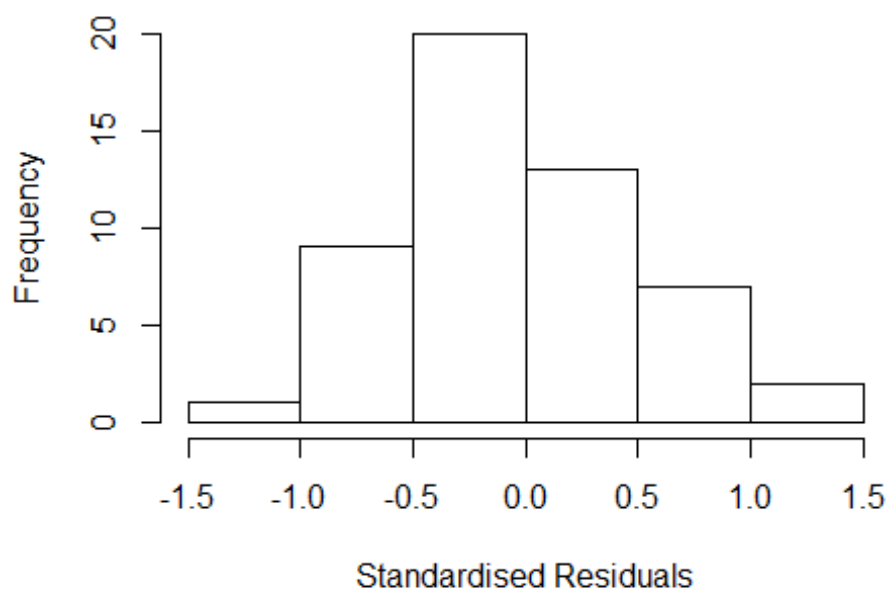
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.11481 -0.30027 -0.00276 0.29250 1.28308
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.05125 0.24276 4.33 0.00013 ***
## FirstAuthorFemale1 0.42047 0.22230 1.89 0.06737 .
## LastAuthorFemale1 -0.30907 0.22115 -1.40 0.17157
## UniqueAuthors2 0.04989 0.23904 0.21 0.83596
## UniqueAuthors3 -0.88573 0.18532 -4.78 3.5e-05 ***
## UniqueAuthors4 0.95304 0.28695 3.32 0.00220 **
## UniqueAuthors5 -0.23924 0.29316 -0.82 0.42031
## Year1999 0.08435 0.19100 0.44 0.66165
## Year2001 -0.32825 0.24276 -1.35 0.18552
## Year2002 -0.17195 0.31094 -0.55 0.58398
```

```

## Year2004          -0.30103      0.40211      -0.75      0.45938
## Year2005          -0.56925      0.24276      -2.34      0.02520 *
## Year2006          -0.07207      0.17915      -0.40      0.69007
## Year2007          -0.49670      0.27594      -1.80      0.08100 .
## Year2008          -0.69325      0.24276      -2.86      0.00737 **
## Year2009          -0.54433      0.43277      -1.26      0.21730
## Year2010          -0.00862      0.36276      -0.02      0.98119
## Year2011           0.17306      0.32309       0.54      0.59579
## Year2012          -0.09773      0.26888      -0.36      0.71858
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.583
## Multiple R-squared:  0.25,    Adjusted R-squared:  -0.159
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.607  0.866  0.971   0.924   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          1.92e-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 7.039e+15  1      8.390e+07
## LastAuthorFemale  5.010e+15  1      7.078e+07
## Year              6.504e+15 12      4.559e+00

```

Residuals from first and last author



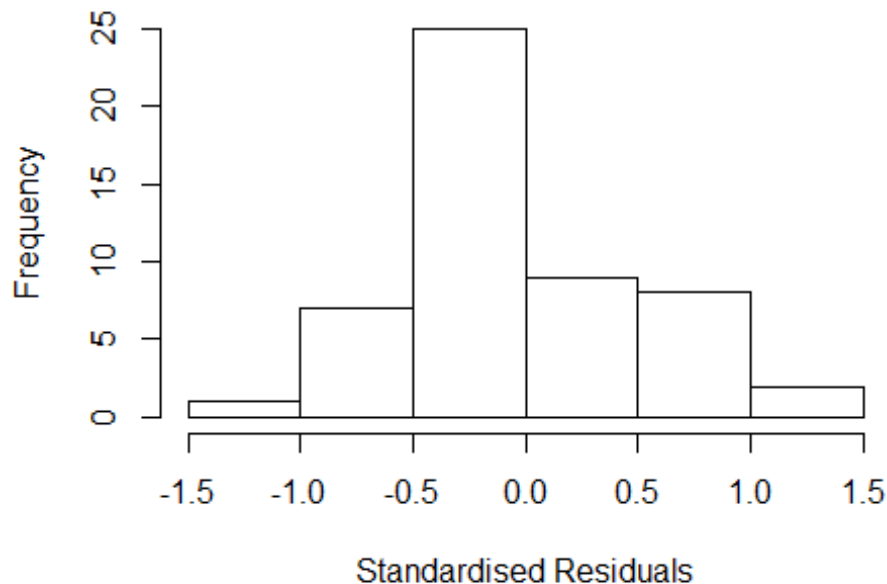
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.07111 -0.32336 -0.00755 0.30590 1.27790
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9313 0.1955 4.76 2.9e-05 ***
## FirstAuthorFemale1 0.3254 0.2603 1.25 0.2191
## LastAuthorFemale1 -0.1682 0.2405 -0.70 0.4887
## Year1999 0.1586 0.2003 0.79 0.4336
## Year2001 -0.2083 0.1955 -1.07 0.2937
## Year2002 -0.0749 0.3138 -0.24 0.8128
## Year2004 -0.2192 0.3738 -0.59 0.5611
## Year2005 -0.4493 0.1955 -2.30 0.0273 *
## Year2006 0.0489 0.1781 0.27 0.7853
## Year2007 -0.0605 0.3426 -0.18 0.8607
## Year2008 -0.5733 0.1955 -2.93 0.0057 **
## Year2009 -0.4192 0.4151 -1.01 0.3191
```

```

## Year2010          0.0343      0.3284      0.10      0.9175
## Year2011          0.2557      0.3241      0.79      0.4350
## Year2012         -0.0173      0.2694     -0.06      0.9491
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.629
## Multiple R-squared:  0.167, Adjusted R-squared:  -0.148
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.659  0.879  0.946  0.923  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      1.92e-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 -2.256e+16  1          NaN
## Year              -2.256e+16 12          NaN

```

Residuals from first author



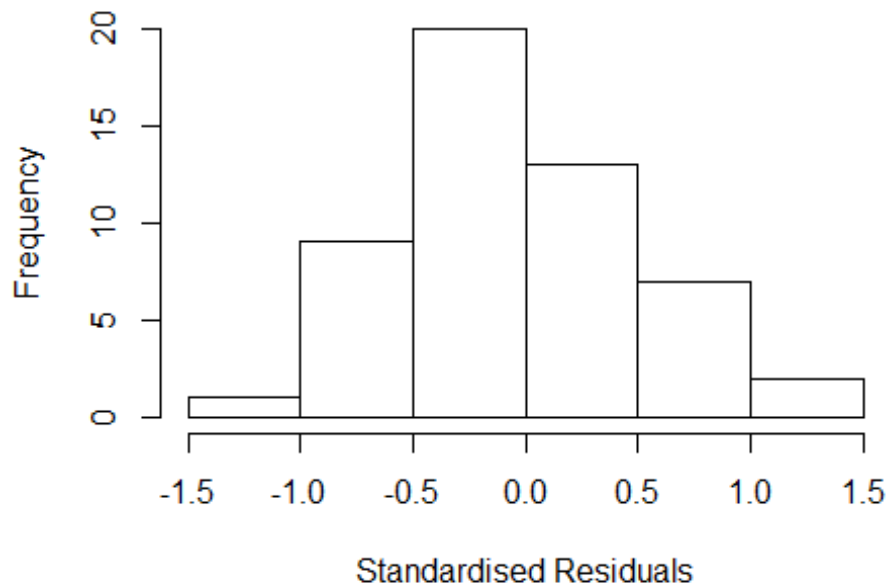
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0714 -0.3278 -0.0353 0.3719 1.2994
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.97537 0.20829 4.68 3.6e-05 ***
## FirstAuthorFemale1 0.18807 0.20114 0.94 0.3557
## Year1999 0.08356 0.19852 0.42 0.6762
## Year2001 -0.25237 0.20829 -1.21 0.2331
## Year2002 -0.13441 0.33071 -0.41 0.6867
## Year2004 -0.28426 0.36287 -0.78 0.4383
## Year2005 -0.49337 0.20829 -2.37 0.0230 *
## Year2006 -0.00295 0.18919 -0.02 0.9876
## Year2007 -0.11411 0.34393 -0.33 0.7419
## Year2008 -0.61737 0.20829 -2.96 0.0052 **
## Year2009 -0.48475 0.41183 -1.18 0.2465
## Year2010 -0.05100 0.29936 -0.17 0.8656
```

```

## Year2011          0.18090    0.32320    0.56    0.5790
## Year2012          -0.09202    0.27173   -0.34    0.7367
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.166, Adjusted R-squared:  -0.12
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.623  0.872  0.943  0.917  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      1.92e-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 5.274e+13 1      7.262e+06
## Year              5.274e+13 12      3.730e+00

```

Residuals from last author



```
## [1] "List of 0 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
## -1.0154 -0.3060 -0.0232 0.3301 1.2681
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0635 0.1817 5.85 9.1e-07 ***
## LastAuthorFemale1 0.1033 0.1951 0.53 0.5995
## Year1999 0.0803 0.2162 0.37 0.7125
## Year2001 -0.3405 0.1817 -1.87 0.0687 .
## Year2002 -0.1801 0.2991 -0.60 0.5507
## Year2004 -0.3098 0.4088 -0.76 0.4532
## Year2005 -0.5815 0.1817 -3.20 0.0028 **
## Year2006 -0.0698 0.1753 -0.40 0.6929
## Year2007 -0.1771 0.3540 -0.50 0.6197
## Year2008 -0.7055 0.1817 -3.88 0.0004 ***
## Year2009 -0.5415 0.4112 -1.32 0.1958
## Year2010 -0.1344 0.3015 -0.45 0.6582
```

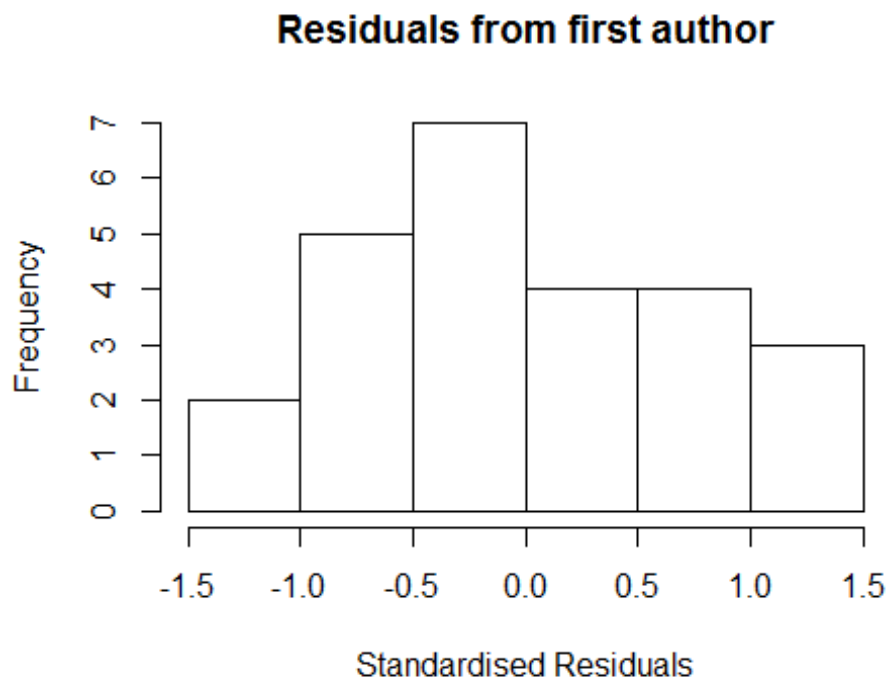
```

## Year2011          0.1708      0.3291      0.52      0.6068
## Year2012          -0.1514      0.2585     -0.59      0.5616
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.608
## Multiple R-squared:  0.151, Adjusted R-squared:  -0.139
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.643  0.875  0.944  0.921  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           1.92e-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: 52"
## [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
##    3    8    4    2    5    4    2    4    1    1    3    3    2    1    1
## 2012
##    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    8    2    0    0    1    2    3    0    1    2    2    2    0    0
## 2012
##    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    7    2    0    0    1    2    3    0    1    2    2    2    0    0

```



```
## 2012
## 1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [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 1.493e+15 1      3.864e+07
## Year              1.493e+15 10      5.737e+00
```



```
## [1] "Regression 4: Last author gender, Year as factors"
##
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.047e+16 1      1.023e+08
## Year              1.047e+16 10      6.324e+00

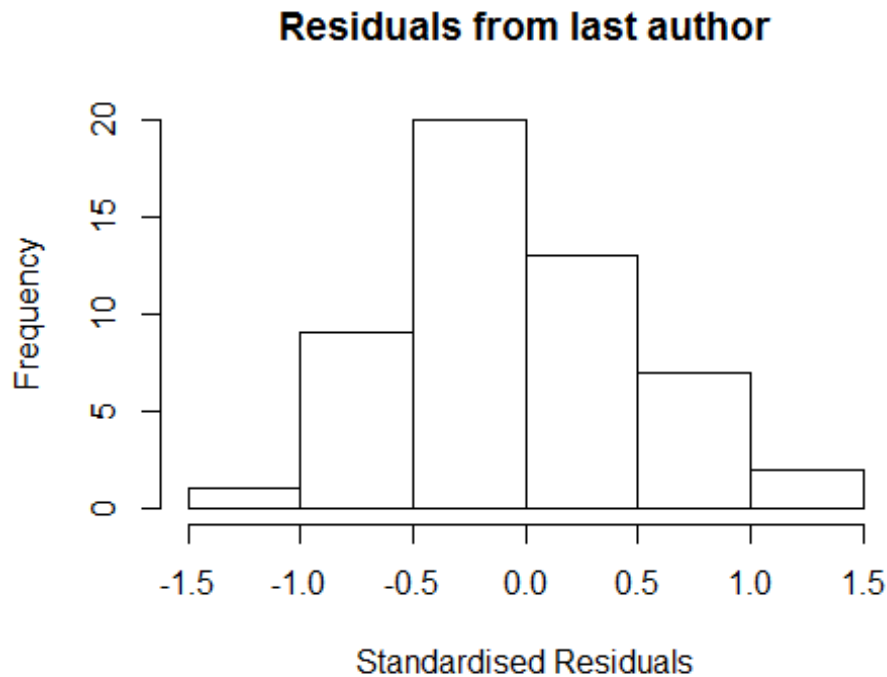
## [1] "Sample size for the above analysis: 25"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2912"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
```

```

gendered] [check that these decrease]"
##
## 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    1    2    4    4    2    2    3    3
##
## 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011 2012
##    0    1    1    1    2    4    4    2    1    3    3
##
## 2001 2002 2003 2004 2006 2007 2008 2009 2010 2011 2012
##    0    1    1    1    2    4    3    2    1    3    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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 20"
## [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]"
##
## 2001 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    2    6    2    1    4    3    2
##
## 2001 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    2    5    2    1    4    3    2
##
## 2001 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    2    5    2    1    4    3    2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.87298334620742"
## [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): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

```

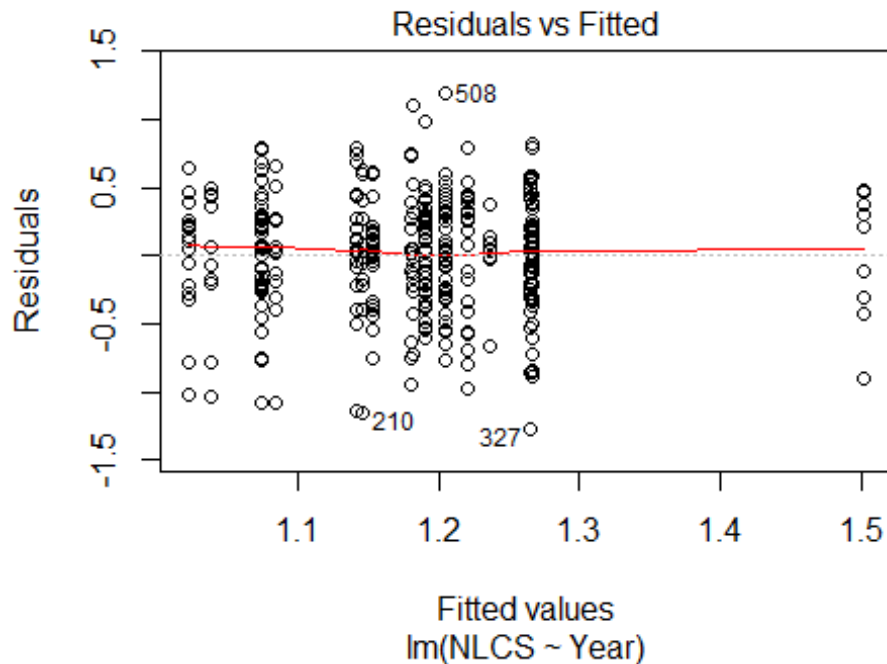


```
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"
## [1] "Regression 3: First author gender, Year as factors"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 20"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2914"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1999 2004 2006 2009 2010 2011 2012
##    1    1    1    2    1    2    3
##
## 1999 2004 2006 2009 2010 2011 2012
##    0    1    1    1    1    2    2
##
## 1999 2004 2006 2009 2010 2011 2012
##    0    1    1    1    1    1    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: 5"
## [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: 7"
## [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
##    21    16    18    19    21    26    32    20    17    22    31    40    28    40    50
## 2011 2012
##    48    59
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    10     8    13    11    12    11    20    13    11    16    23    26    24    30    43
## 2011 2012
##    42    47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    10     8    13    10     8    11    17    11     9    13    23    24    22    28    39
## 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 = 15, df = 16, p-value = 0.5

```



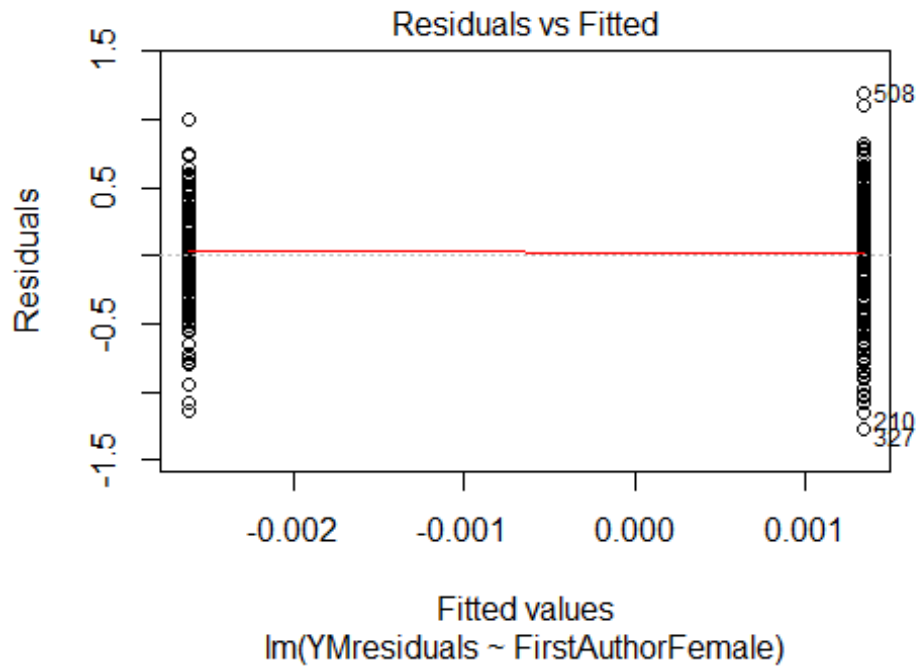
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.023, df = 1, p-value = 0.9

## [1] "Female first author team size 2018 geometric mean: 5.71287273428289"
## [1] "Male first author team size 2018 geometric mean: 3.80931712081705"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

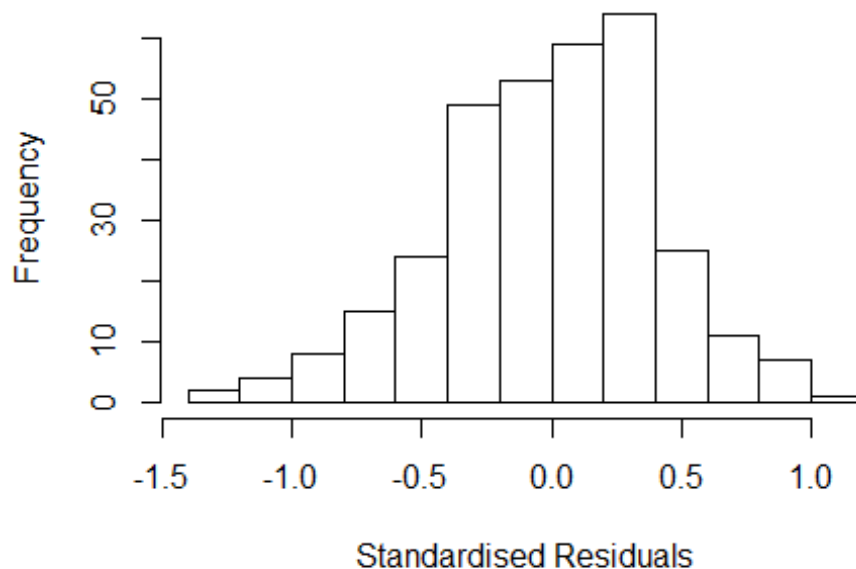
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 300, p-value = 0.02
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.96600776429401"
## [1] "Male last author team size 2018 geometric mean: 5.07312654770042"

## 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 = 230, 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 1.329 1          1.153
## LastAuthorFemale  1.511 1          1.229
## UniqueAuthors     4.237 4          1.198
## Year              4.978 16          1.051
```

Residuals from first and last author and team size



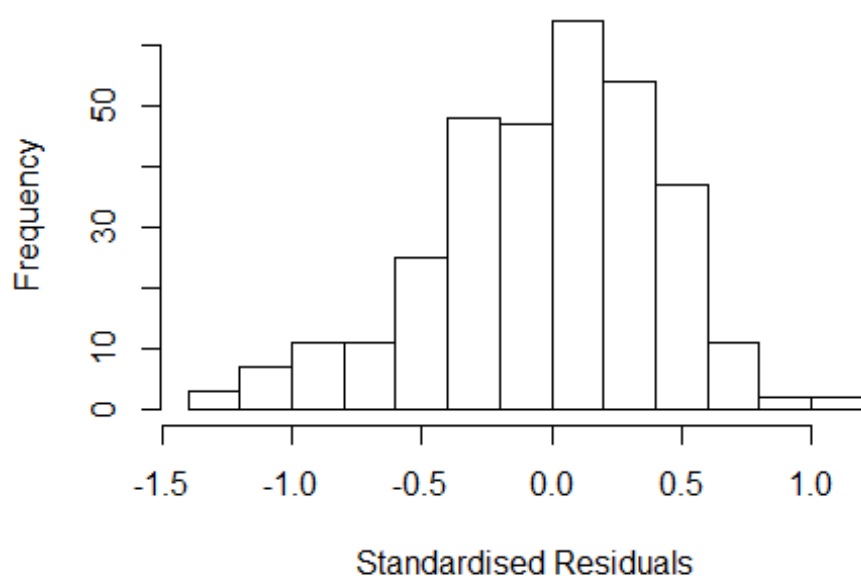
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3225 -0.3200 0.0269 0.2891 1.1994
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1784 0.1978 5.96 7.2e-09 ***
## FirstAuthorFemale1 -0.0301 0.0533 -0.57 0.5722
## LastAuthorFemale1 -0.0105 0.0561 -0.19 0.8511
## UniqueAuthors2 0.1760 0.1742 1.01 0.3129
## UniqueAuthors3 0.3388 0.1541 2.20 0.0287 *
## UniqueAuthors4 0.4341 0.1598 2.72 0.0070 **
## UniqueAuthors5 0.3427 0.1615 2.12 0.0347 *
## Year1997 -0.2030 0.1522 -1.33 0.1833
## Year1998 -0.1816 0.1820 -1.00 0.3191
## Year1999 -0.2326 0.2122 -1.10 0.2740
```

```

## Year2000          -0.0980      0.2251   -0.44   0.6637
## Year2001          -0.3207      0.2062   -1.56   0.1210
## Year2002          -0.1881      0.1980   -0.95   0.3428
## Year2003          -0.4354      0.1901   -2.29   0.0227 *
## Year2004          -0.1814      0.2528   -0.72   0.4737
## Year2005          -0.4586      0.1711   -2.68   0.0078 **
## Year2006          -0.2905      0.1520   -1.91   0.0570 .
## Year2007          -0.1734      0.1689   -1.03   0.3056
## Year2008          -0.2562      0.1677   -1.53   0.1278
## Year2009          -0.1942      0.1542   -1.26   0.2089
## Year2010          -0.2789      0.1485   -1.88   0.0612 .
## Year2011          -0.2730      0.1498   -1.82   0.0694 .
## Year2012          -0.3338      0.1508   -2.21   0.0277 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.107, Adjusted R-squared:  0.0412
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 296 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.286  0.886  0.946  0.903  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          3.11e-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 1.202 1          1.096
## LastAuthorFemale  1.393 1          1.180
## Year              1.616 16          1.015

```


Residuals from first and last author



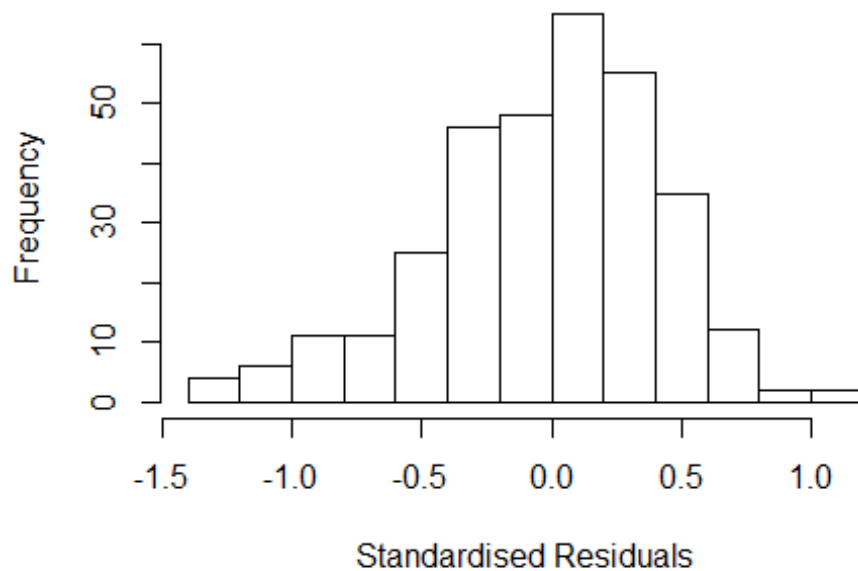
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3131 -0.2853  0.0193  0.2858  1.1787
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5270     0.1428   10.70  <2e-16 ***
## FirstAuthorFemale1  0.0169     0.0518    0.33   0.744
## LastAuthorFemale1 -0.0147     0.0551   -0.27   0.790
## Year1997          -0.2806     0.1681   -1.67   0.096 .
## Year1998          -0.2524     0.1806   -1.40   0.163
## Year1999          -0.3910     0.2388   -1.64   0.103
## Year2000          -0.3224     0.2091   -1.54   0.124
## Year2001          -0.4160     0.2164   -1.92   0.055 .
## Year2002          -0.3131     0.2103   -1.49   0.138
## Year2003          -0.4286     0.2041   -2.10   0.037 *
## Year2004          -0.2911     0.2232   -1.30   0.193
## Year2005          -0.4830     0.1870   -2.58   0.010 *
```

```

## Year2006          -0.3722      0.1616    -2.30      0.022 *
## Year2007          -0.2309      0.1738    -1.33      0.185
## Year2008          -0.3205      0.1880    -1.70      0.089 .
## Year2009          -0.2433      0.1636    -1.49      0.138
## Year2010          -0.3110      0.1593    -1.95      0.052 .
## Year2011          -0.3203      0.1583    -2.02      0.044 *
## Year2012          -0.4001      0.1583    -2.53      0.012 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.423
## Multiple R-squared:  0.0403, Adjusted R-squared:  -0.0167
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 291 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.314  0.878  0.947  0.900  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      3.11e-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 1.205 1      1.098
## Year              1.205 16      1.006

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3064 -0.2875 0.0198 0.2860 1.1722
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5266 0.1439 10.61 <2e-16 ***
## FirstAuthorFemale1 0.0138 0.0522 0.26 0.7920
## Year1997 -0.2795 0.1694 -1.65 0.0999 .
## Year1998 -0.2554 0.1802 -1.42 0.1575
## Year1999 -0.3959 0.2368 -1.67 0.0956 .
## Year2000 -0.3246 0.2103 -1.54 0.1237
## Year2001 -0.4253 0.2116 -2.01 0.0454 *
## Year2002 -0.3224 0.2024 -1.59 0.1122
## Year2003 -0.4387 0.1977 -2.22 0.0272 *
## Year2004 -0.2977 0.2242 -1.33 0.1852
## Year2005 -0.4861 0.1870 -2.60 0.0098 **
## Year2006 -0.3765 0.1603 -2.35 0.0195 *
```

```

## Year2007          -0.2340      0.1732   -1.35    0.1778
## Year2008          -0.3238      0.1874   -1.73    0.0850 .
## Year2009          -0.2489      0.1616   -1.54    0.1245
## Year2010          -0.3156      0.1590   -1.98    0.0481 *
## Year2011          -0.3239      0.1584   -2.04    0.0417 *
## Year2012          -0.4043      0.1579   -2.56    0.0109 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.424
## Multiple R-squared:  0.0401, Adjusted R-squared:  -0.0136
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 34 weights are ~= 1. The remaining 288 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.321  0.878  0.947  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      3.11e-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 1.391 1      1.179
## Year      1.391 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
## -1.306 -0.293  0.022  0.277  1.183

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5345    0.1433   10.71 <2e-16 ***
## LastAuthorFemale1 -0.0113    0.0552   -0.20  0.838
## Year1997          -0.2805    0.1692   -1.66  0.098 .
## Year1998          -0.2508    0.1821   -1.38  0.169
## Year1999          -0.3914    0.2403   -1.63  0.104
## Year2000          -0.3251    0.2106   -1.54  0.124
## Year2001          -0.4186    0.2178   -1.92  0.056 .
## Year2002          -0.3111    0.2114   -1.47  0.142
## Year2003          -0.4281    0.2049   -2.09  0.037 *
## Year2004          -0.2889    0.2233   -1.29  0.197
## Year2005          -0.4794    0.1884   -2.54  0.011 *
## Year2006          -0.3694    0.1624   -2.27  0.024 *
## Year2007          -0.2287    0.1744   -1.31  0.191
## Year2008          -0.3150    0.1878   -1.68  0.095 .
## Year2009          -0.2412    0.1647   -1.46  0.144
## Year2010          -0.3092    0.1598   -1.93  0.054 .
## Year2011          -0.3180    0.1587   -2.00  0.046 *
## Year2012          -0.3973    0.1594   -2.49  0.013 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.421
## Multiple R-squared:  0.04, Adjusted R-squared:  -0.0137
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 292 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.316  0.875   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      3.11e-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: 322"
## [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]"
##
## 2009 2010
##    1    2
##
## 2009 2010
##    1    2
##
## 2009 2010
##    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2"
## [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 2919"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1998 1999 2001 2002 2004 2005 2006 2008 2009 2010 2011 2012
##    1    2    2    1    1    1    3    1    1    6    1    4    4
##
## 1996 1998 1999 2001 2002 2004 2005 2006 2008 2009 2010 2011 2012
##    1    1    1    0    1    1    3    1    1    6    1    4    4
##
## 1996 1998 1999 2001 2002 2004 2005 2006 2008 2009 2010 2011 2012
##    1    1    1    0    1    1    3    1    1    6    1    4    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [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"

```

```

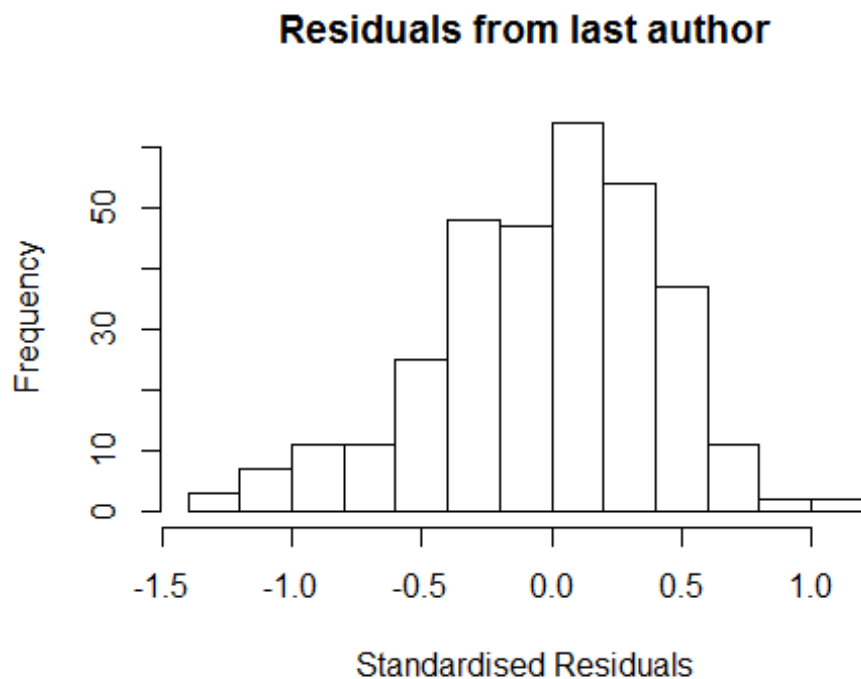
## [1] "Regression 3: First author gender, Year as factors"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 25"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 2920"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 2002 2009
##    1    1
##
## 2002 2009
##    1    1
##
## 2002 2009
##    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5"
## [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 2921"
## [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
##    3    2    2    2    6    6    1    7    4    6    4   12    8    5    9
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    0    0    1    6    4    1    6    2    6    3   10    7    3    5
##
## 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    0    0    1    6    3    1    6    1    6    3   10    7    3    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.4456890899877"
## [1] "Male first author team size 2018 geometric mean: 2.0800838230519"

```

```
## 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 = 14, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.21336383940064"
## [1] "Male last author team size 2018 geometric mean: 2.41565251562535"

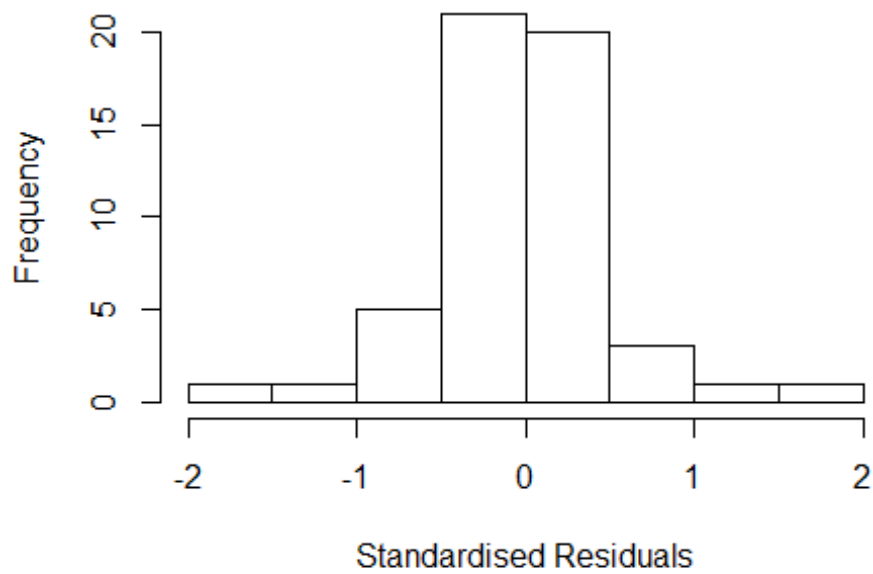
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 13, 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 4.680e+14  1      2.163e+07
```


## LastAuthorFemale	1.892e+14	1	1.375e+07
## UniqueAuthors	2.041e+17	4	1.458e+02
## Year	2.253e+31	12	2.025e+01

Residuals from first and last author and team size



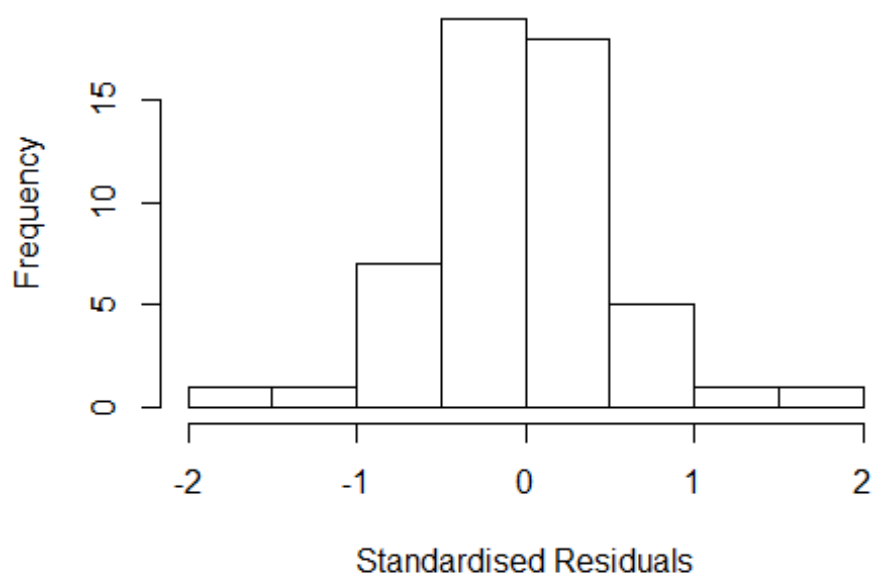
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.73e+00 -2.86e-01 -4.44e-16 2.84e-01 1.56e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9280 0.2152 4.31 0.00013 ***
## FirstAuthorFemale1 0.3294 0.1902 1.73 0.09228 .
## LastAuthorFemale1 0.2800 0.2346 1.19 0.24085
## UniqueAuthors2 0.2128 0.3064 0.69 0.49213
## UniqueAuthors3 0.2895 0.3407 0.85 0.40147
## UniqueAuthors4 0.7497 0.3091 2.43 0.02075 *
```

```

## UniqueAuthors5      0.3787      0.2977      1.27  0.21202
## Year2001             -0.3140      0.2152     -1.46  0.15377
## Year2002             -0.3750      0.4791     -0.78  0.43924
## Year2003             -0.6524      0.3187     -2.05  0.04844 *
## Year2004             -0.3234      0.3046     -1.06  0.29584
## Year2005             -0.0237      0.3531     -0.07  0.94689
## Year2006             -0.9647      0.4258     -2.27  0.02996 *
## Year2007             -0.6519      0.5821     -1.12  0.27060
## Year2008             -0.2326      0.3911     -0.59  0.55600
## Year2009             -0.5081      0.3467     -1.47  0.15206
## Year2010             -0.4472      0.3215     -1.39  0.17329
## Year2011             -0.9588      0.4233     -2.26  0.03000 *
## Year2012             -0.4580      0.3456     -1.33  0.19394
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.501
## Multiple R-squared:  0.418, Adjusted R-squared:  0.11
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.210  0.915  0.967  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      1.89e-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 1.511e+14 1      1.229e+07
## LastAuthorFemale  2.907e+14 1      1.705e+07
## Year              2.785e+28 12      1.532e+01

```

Residuals from first and last author



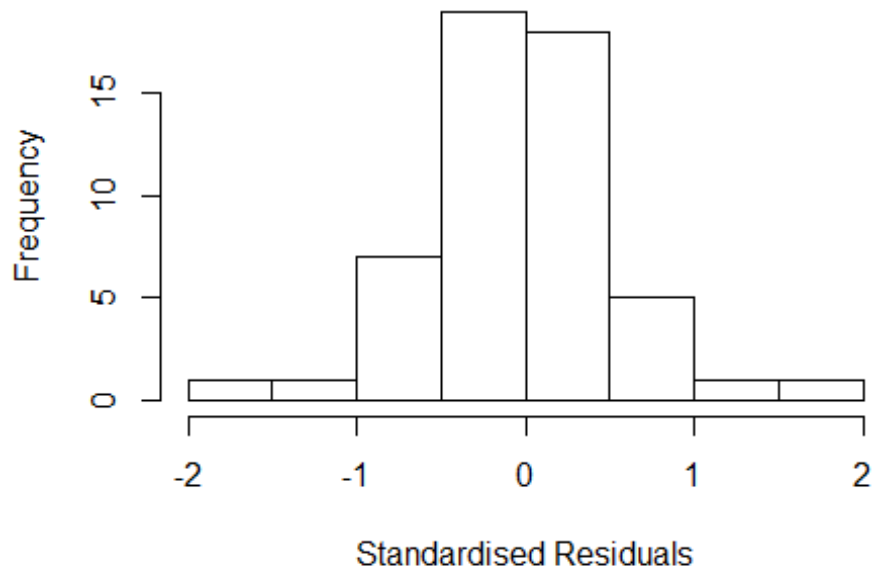
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.69e+00 -2.92e-01 -2.08e-16 2.84e-01 1.62e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.92800 0.21058 4.41 8.3e-05 ***
## FirstAuthorFemale1 0.43264 0.19320 2.24 0.031 *
## LastAuthorFemale1 0.17015 0.19876 0.86 0.397
## Year2001 -0.31400 0.21058 -1.49 0.144
## Year2002 -0.21592 0.36799 -0.59 0.561
## Year2003 -0.44882 0.31649 -1.42 0.164
## Year2004 -0.31679 0.35101 -0.90 0.372
## Year2005 0.15832 0.43247 0.37 0.716
## Year2006 -0.10515 0.28956 -0.36 0.719
## Year2007 -0.32005 0.32158 -1.00 0.326
## Year2008 -0.00793 0.26883 -0.03 0.977
## Year2009 -0.28014 0.35284 -0.79 0.432
```

```

## Year2010          -0.20437    0.29023   -0.70    0.486
## Year2011          -0.49873    0.43704   -1.14    0.261
## Year2012          -0.16198    0.34447   -0.47    0.641
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.616
## Multiple R-squared:  0.26,   Adjusted R-squared:  -0.0132
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 45 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.433  0.882  0.973  0.921  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      1.89e-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 -3.031e+15  1      NaN
## Year              -3.031e+15 12      NaN

```

Residuals from first author



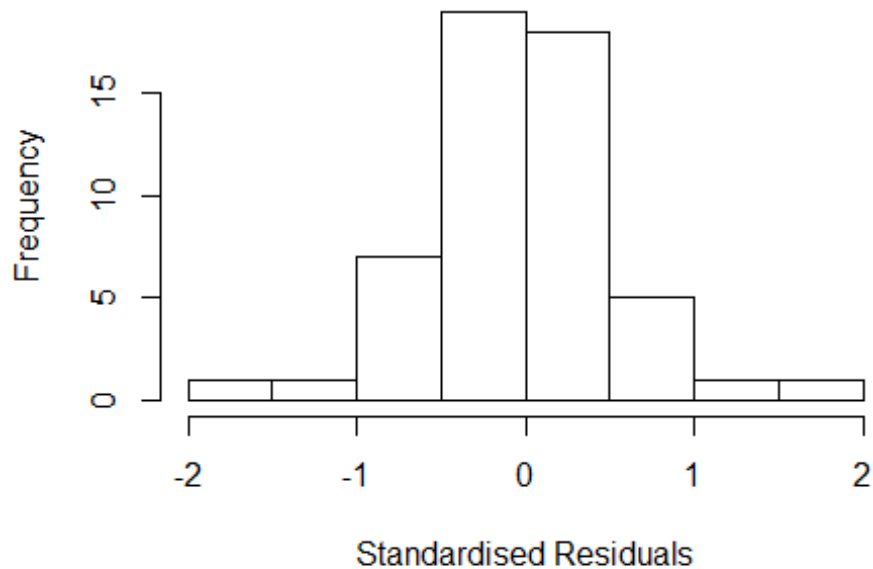
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.65e+00 -2.83e-01  1.11e-16  2.86e-01  1.57e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9280     0.2110   4.40 8.2e-05 ***
## FirstAuthorFemale1 0.4748     0.2130   2.23  0.032 *
## Year2001        -0.3140     0.2110  -1.49  0.145
## Year2002        -0.2311     0.3698  -0.62  0.536
## Year2003        -0.4086     0.3485  -1.17  0.248
## Year2004        -0.1888     0.2998  -0.63  0.533
## Year2005         0.2441     0.4018   0.61  0.547
## Year2006         0.0650     0.2110   0.31  0.760
## Year2007        -0.3361     0.3234  -1.04  0.305
## Year2008        -0.0357     0.2730  -0.13  0.897
## Year2009        -0.2317     0.3292  -0.70  0.486
## Year2010        -0.1529     0.2893  -0.53  0.600
```

```

## Year2011          -0.3702      0.3988   -0.93    0.359
## Year2012          -0.0785      0.3270   -0.24    0.812
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.251, Adjusted R-squared:  0.00177
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.437  0.882  0.957  0.909  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      1.89e-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 -3234 1      NaN
## Year             -3234 12      NaN

```

Residuals from last author



```
## [1] "List of 0 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
## -1.52e+00 -3.77e-01 -2.78e-16 4.22e-01 1.45e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.92800 0.20771 4.47 6.6e-05 ***
## LastAuthorFemale1 0.28736 0.23571 1.22 0.23
## Year2001 -0.31400 0.20771 -1.51 0.14
## Year2002 -0.07016 0.39298 -0.18 0.86
## Year2003 -0.34357 0.35595 -0.97 0.34
## Year2004 -0.00136 0.31417 0.00 1.00
## Year2005 0.30901 0.39654 0.78 0.44
## Year2006 -0.22236 0.31417 -0.71 0.48
## Year2007 -0.15800 0.27028 -0.58 0.56
## Year2008 0.28047 0.28492 0.98 0.33
## Year2009 -0.11153 0.33531 -0.33 0.74
## Year2010 -0.07156 0.32569 -0.22 0.83
```

```

## Year2011          -0.18728    0.41253   -0.45    0.65
## Year2012          0.18267    0.30368    0.60    0.55
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.755
## Multiple R-squared:  0.149, Adjusted R-squared:  -0.134
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.663  0.941  0.968  0.943  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      1.89e-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: 53"
## [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]"
##
## 2004 2006 2007 2008 2009 2010 2011 2012
##    1    3    2    2    1    1    3    3
##
## 2004 2006 2007 2008 2009 2010 2011 2012
##    1    2    2    2    1    1    3    3
##
## 2004 2006 2007 2008 2009 2010 2011 2012
##    1    2    2    2    1    1    3    3
## [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"

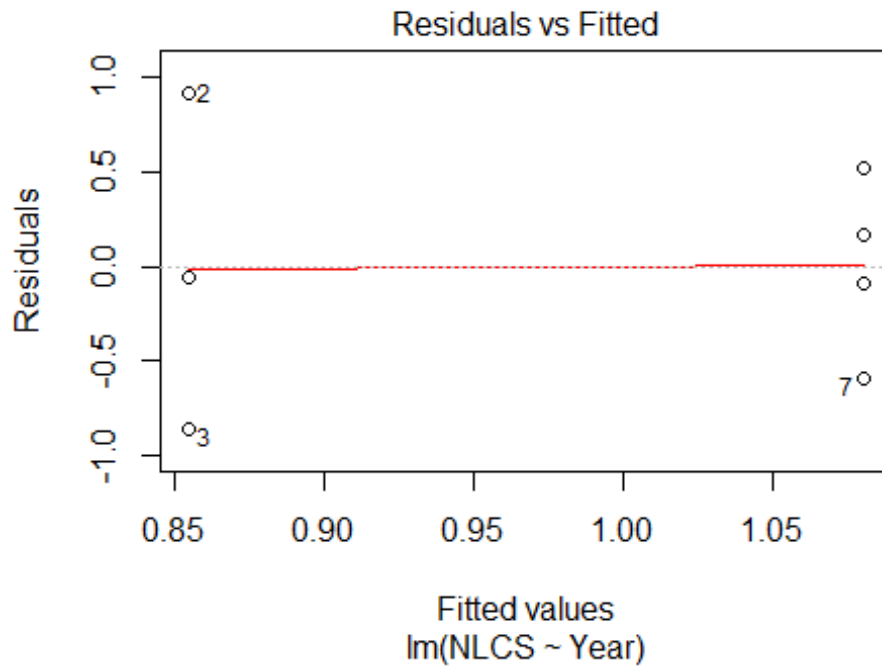
```



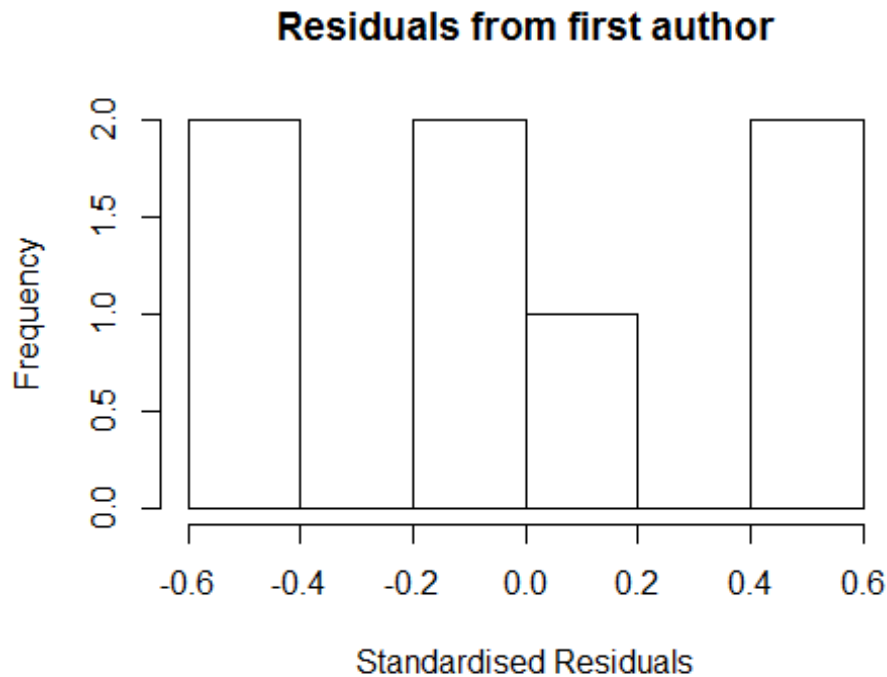
```

## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 15"
## [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]"
##
## 2011 2012
##    3    4
##
## 2011 2012
##    3    4
##
## 2011 2012
##    3    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 0.83, df = 1, p-value = 0.4

```



```
## [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"
## FirstAuthorFemale      Year
##           4.178           4.178
```



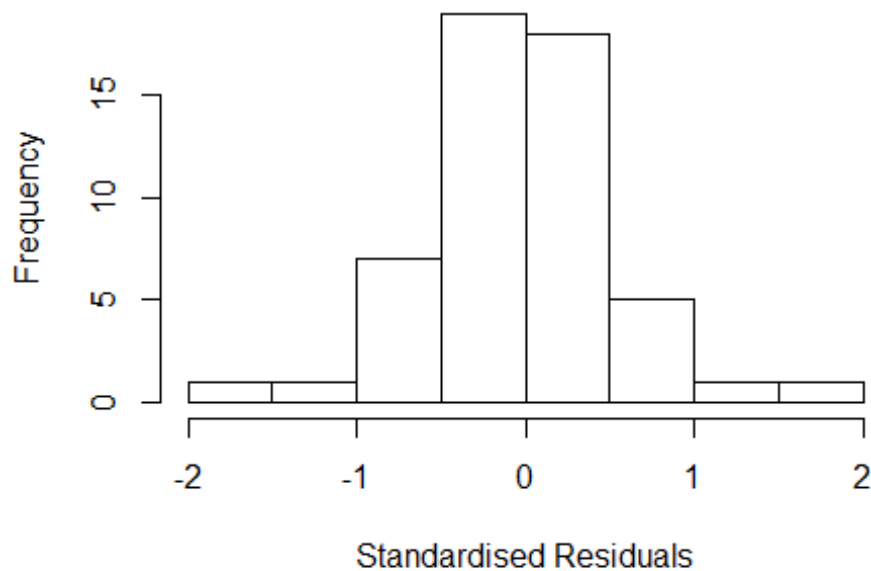
```
## [1] "Regression 4: Last author gender, Year as factors"
## LastAuthorFemale      Year
##           2.299           2.299

## [1] "Sample size for the above analysis: 7"
## [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
##    3    2    5    7    3    5    3    6   10   12    3    5    2    2    1
## 2011 2012
##    3    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    2    5    2    2    1    5   10   12    3    4    2    1    1
## 2011 2012
##    2    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    1    5    2    2    1    5   10   12    3    3    2    1    0
## 2011 2012
##    2    1
```

```
## [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: 3.46410161513775"

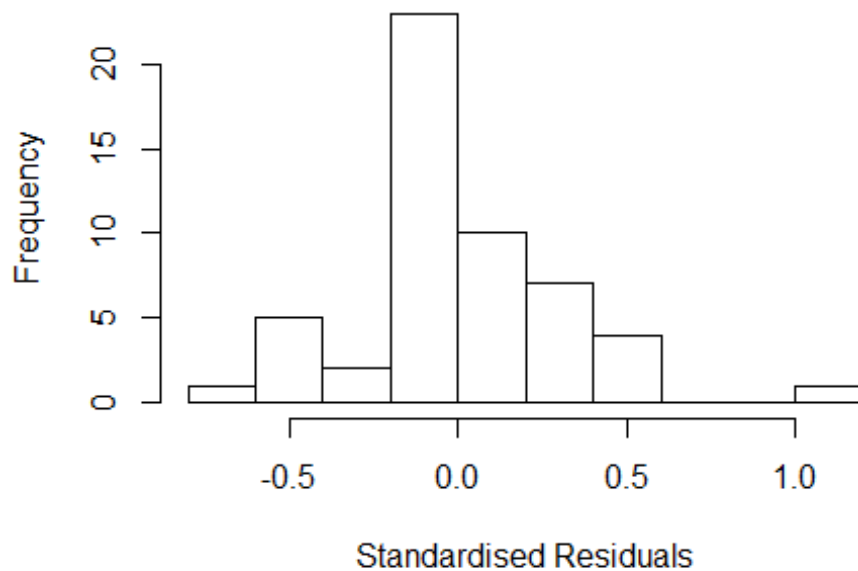
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 3.63424118566428"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 3.211e+14 1      17919235
## LastAuthorFemale -1.544e+15 1          NaN
## UniqueAuthors    -9.679e+43 4          NaN
## Year             -5.113e+45 15          NaN
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -6.83e-01 -1.42e-01 -1.11e-16 1.57e-01 1.08e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.06e+00 1.85e-01 1.12e+01 2.2e-12 ***
## FirstAuthorFemale1 -3.00e-01 2.02e-01 -1.49e+00 0.14744
## LastAuthorFemale1 8.92e-02 2.03e-01 4.40e-01 0.66356
## UniqueAuthors2 1.89e-02 2.05e-01 9.00e-02 0.92703
## UniqueAuthors3 3.65e-01 2.11e-01 1.73e+00 0.09394 .
## UniqueAuthors4 1.03e+00 1.68e-01 6.10e+00 9.1e-07 ***
## UniqueAuthors5 6.02e-01 1.93e-01 3.12e+00 0.00394 **
## Year1997 -1.38e+00 2.47e-01 -5.59e+00 4.0e-06 ***
## Year1998 -9.30e-01 1.85e-01 -5.04e+00 1.9e-05 ***
## Year1999 -1.04e+00 2.98e-01 -3.50e+00 0.00143 **
```

```

## Year2000      -1.42e+00    2.64e-01 -5.38e+00    7.3e-06 ***
## Year2001      -1.20e+00    1.00e-01 -1.20e+01    3.4e-13 ***
## Year2002      -7.96e-01    2.70e-08 -2.95e+07    < 2e-16 ***
## Year2003      -1.16e+00    3.07e-01 -3.80e+00    0.00064 ***
## Year2004      -1.23e+00    2.32e-01 -5.29e+00    9.3e-06 ***
## Year2005      -1.36e+00    2.72e-01 -4.99e+00    2.2e-05 ***
## Year2006      -1.37e+00    2.10e-01 -6.53e+00    2.7e-07 ***
## Year2007      -1.98e+00    2.22e-01 -8.95e+00    4.2e-10 ***
## Year2008      -6.16e-01    1.43e-01 -4.30e+00    0.00016 ***
## Year2009      -2.09e+00    3.10e-01 -6.75e+00    1.5e-07 ***
## Year2011      -1.42e+00    4.61e-01 -3.08e+00    0.00437 **
## Year2012      -1.64e+00    3.06e-01 -5.35e+00    7.8e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.337
## Multiple R-squared:  0.707, Adjusted R-squared:  0.509
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.286  0.881  0.974  0.922  0.986  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      1.89e-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 -4.483e+15  1      NaN
## LastAuthorFemale -2.077e+15  1      NaN
## Year      -1.881e+30 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 + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q      Median        3Q       Max
## -1.05e+00 -2.33e-01 -3.33e-16  2.65e-01  1.04e+00
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    2.17e+00   2.21e-01   9.84e+00  1.3e-11 ***
## FirstAuthorFemale1 -3.93e-01   2.21e-01  -1.78e+00  0.08329 .
## LastAuthorFemale1  -3.91e-01   1.84e-01  -2.13e+00  0.04059 *
## Year1997         -1.31e+00   2.21e-01  -5.92e+00  9.7e-07 ***
## Year1998         -1.04e+00   2.21e-01  -4.72e+00  3.7e-05 ***
## Year1999         -5.76e-01   4.05e-01  -1.42e+00  0.16377
## Year2000         -8.37e-01   4.89e-01  -1.71e+00  0.09603 .
## Year2001         -1.26e+00   1.10e-01  -1.14e+01  2.3e-13 ***
## Year2002         -7.96e-01   2.12e-08  -3.75e+07  < 2e-16 ***
## Year2003         -5.43e-01   4.31e-01  -1.26e+00  0.21615
## Year2004         -7.15e-01   1.85e-01  -3.87e+00  0.00046 ***
## Year2005         -8.59e-01   1.84e-01  -4.67e+00  4.4e-05 ***
## Year2006         -7.34e-01   1.97e-01  -3.72e+00  0.00069 ***
## Year2007         -1.77e+00   3.21e-01  -5.52e+00  3.3e-06 ***
## Year2008         -3.75e-01   3.32e-01  -1.13e+00  0.26556
## Year2009         -6.07e-01   1.84e-01  -3.29e+00  0.00226 **
## Year2011         -9.40e-01   3.70e-01  -2.54e+00  0.01567 *
## Year2012         -8.10e-01   1.84e-01  -4.40e+00  9.7e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.428
## Multiple R-squared:  0.497, Adjusted R-squared:  0.252
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 42 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.530  0.896  0.949  0.915  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      1.89e-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

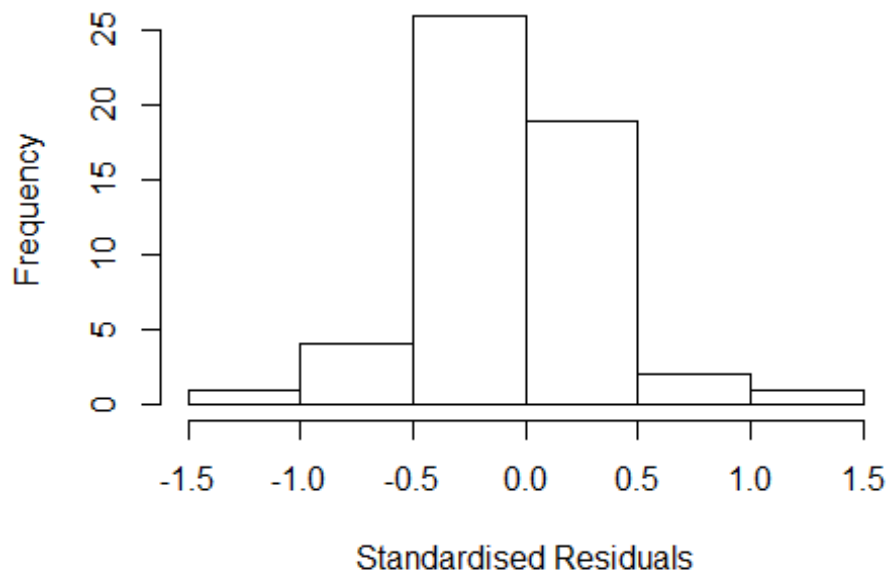
```

```
##           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"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

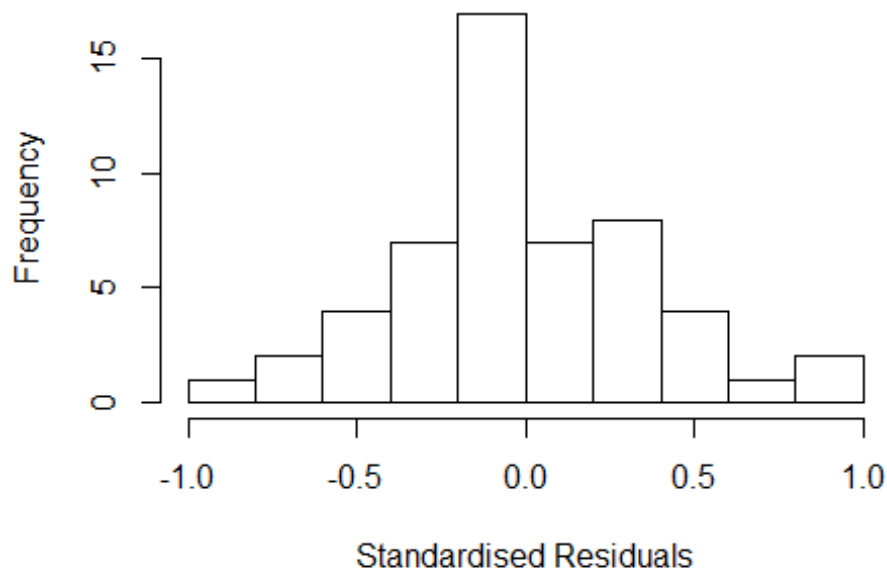
## Warning in cov2cor(v): diag(.) had 0 or NA entries; non-finite result is
## doubtful
```

Residuals from first and last author



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1 NaN
## Year NaN 15 NaN
```


Residuals from first author



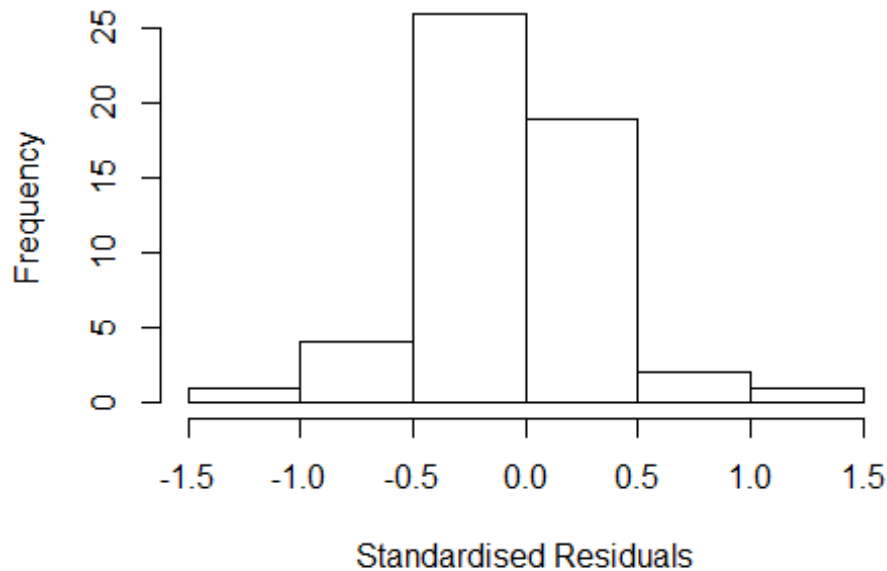
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -9.13e-01 -2.53e-01 1.11e-16 2.77e-01 9.99e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 2.4426 0.2465 9.91 7.9e-12 ***
## FirstAuthorFemale1 -0.6656 0.2465 -2.70 0.01048 *
## Year1997 -1.5836 0.2471 -6.41 2.0e-07 ***
## Year1998 -1.3146 0.2465 -5.33 5.4e-06 ***
## Year1999 -0.9805 0.3856 -2.54 0.01545 *
## Year2000 -1.1091 0.4963 -2.23 0.03174 *
## Year2001 -1.3963 0.1565 -8.92 1.2e-10 ***
## Year2002 -0.7960 0.0000 -Inf < 2e-16 ***
## Year2003 -0.9834 0.3790 -2.59 0.01360 *
## Year2004 -1.0859 0.0766 -14.17 2.7e-16 ***
## Year2005 -1.2056 0.1390 -8.68 2.4e-10 ***
## Year2006 -1.0166 0.2646 -3.84 0.00048 ***
```

```

## Year2007          -2.0467      0.3439   -5.95   8.1e-07 ***
## Year2008          -0.5710      0.1339   -4.27   0.00014 ***
## Year2009          -0.9980      0.0000   -Inf    < 2e-16 ***
## Year2011          -1.2723      0.2934   -4.34   0.00011 ***
## Year2012          -1.2010      0.0000   -Inf    < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.43
## Multiple R-squared:  0.446, Adjusted R-squared:  0.2
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 45 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.569  0.889  0.959  0.916  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      1.89e-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 -1.085e+15  1      NaN
## Year             -1.085e+15 15      NaN

```

Residuals from last author



```
## [1] "List of 0 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
## -1.0289 -0.2266 -0.0249 0.2963 1.0604
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.78e+00 3.03e-08 5.86e+07 < 2e-16 ***
## LastAuthorFemale1 -5.86e-01 1.95e-01 -3.01e+00 0.0048 **
## Year1997 -9.18e-01 1.84e-02 -4.99e+01 < 2e-16 ***
## Year1998 -6.49e-01 2.46e-08 -2.64e+07 < 2e-16 ***
## Year1999 -1.99e-01 3.47e-01 -5.70e-01 0.5693
## Year2000 -4.44e-01 4.25e-01 -1.04e+00 0.3034
## Year2001 -1.06e+00 1.43e-01 -7.42e+00 9.2e-09 ***
## Year2002 -7.96e-01 2.23e-08 -3.57e+07 < 2e-16 ***
## Year2003 -1.77e-01 3.14e-01 -5.60e-01 0.5767
## Year2004 -4.87e-01 2.12e-01 -2.30e+00 0.0275 *
## Year2005 -6.82e-01 2.18e-01 -3.12e+00 0.0035 **
## Year2006 -6.02e-01 1.54e-01 -3.90e+00 0.0004 ***
```

```

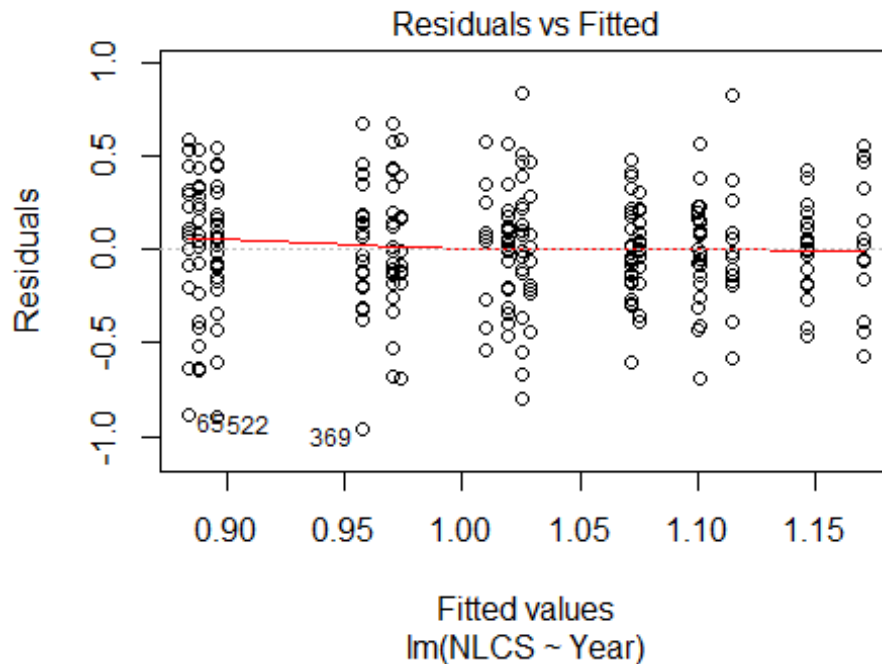
## Year2007          -1.38e+00   2.31e-01 -5.98e+00   7.4e-07 ***
## Year2008          -2.78e-01   4.39e-01 -6.30e-01    0.5299
## Year2009          -4.12e-01   1.95e-01 -2.12e+00    0.0411 *
## Year2011          -6.47e-01   2.46e-01 -2.63e+00    0.0126 *
## Year2012          -6.15e-01   1.95e-01 -3.16e+00    0.0032 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.448
## Multiple R-squared:  0.462, Adjusted R-squared:  0.222
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 44 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.555  0.900   0.958   0.921   0.987   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      1.89e-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: 53"
## [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]"
##
## 2004 2010 2012
##    1    1    1
##
## 2004 2010 2012
##    1    1    1
##
## 2004 2010 2012
##    1    1    1
## [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: 3"
## [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: 3"
## [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
## 12 20 19 26 26 24 15 19 14 23 23 24 24 26 15
## 2011 2012
## 28 40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 14 15 20 19 13 10 14 11 23 20 19 20 19 11
## 2011 2012
## 22 33
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 13 15 18 17 12 10 13 10 22 19 16 19 17 9
## 2011 2012
## 20 28
## [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

```



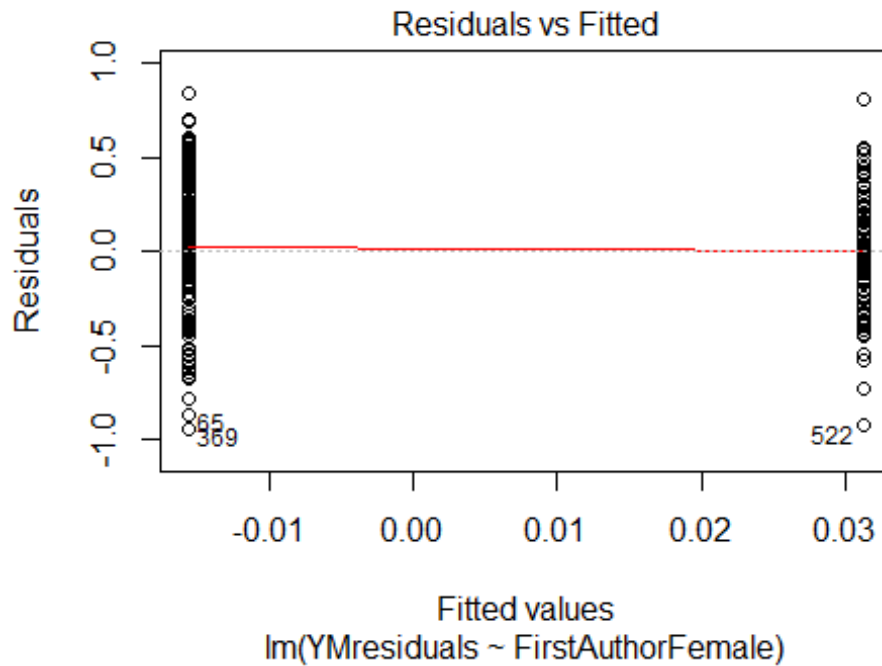
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 5.15850376831545"
## [1] "Male first author team size 2018 geometric mean: 4.11559713783608"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

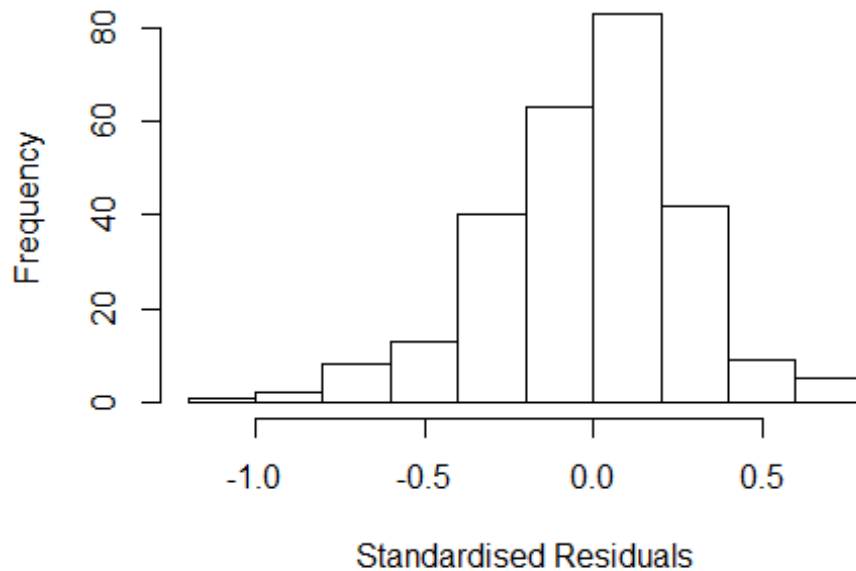
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 35, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.04282321707986"
## [1] "Male last author team size 2018 geometric mean: 4.95484734035914"

## 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 = 22, 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 1.480 1          1.217
## LastAuthorFemale  1.865 1          1.366
## UniqueAuthors    5.389 4          1.234
## Year              8.859 16         1.071
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.0618 -0.1930 0.0188 0.1659 0.7617
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.125676 0.109232 10.31 <2e-16 ***
## FirstAuthorFemale1 0.051174 0.040095 1.28 0.203
## LastAuthorFemale1 0.062939 0.055932 1.13 0.262
## UniqueAuthors2 -0.254426 0.130035 -1.96 0.052 .
## UniqueAuthors3 -0.149969 0.119722 -1.25 0.212
## UniqueAuthors4 -0.094417 0.121311 -0.78 0.437
## UniqueAuthors5 0.088135 0.115887 0.76 0.448
## Year1997 0.092119 0.123867 0.74 0.458
## Year1998 -0.063871 0.135909 -0.47 0.639
## Year1999 0.020560 0.108181 0.19 0.849
```

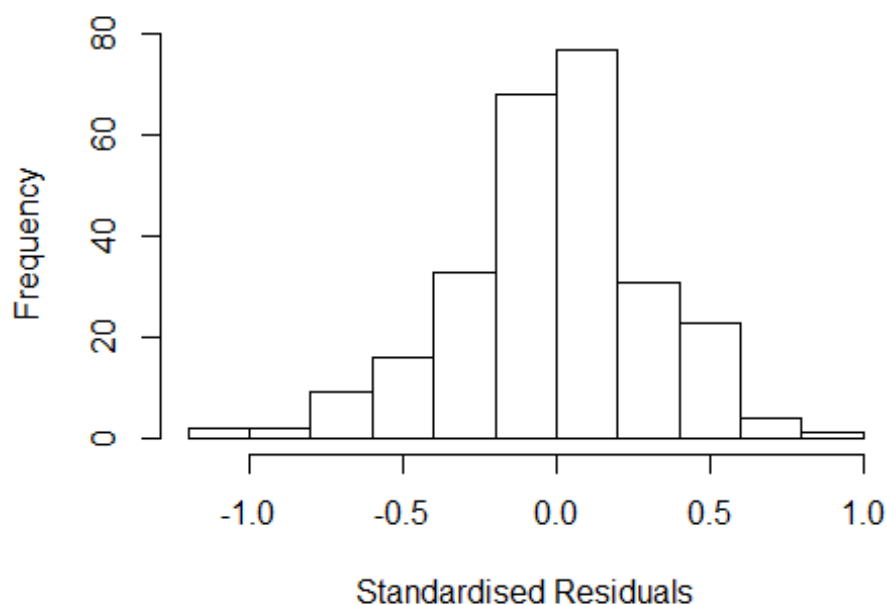


```

## Year2000      0.007301  0.136142  0.05  0.957
## Year2001      0.083544  0.129841  0.64  0.521
## Year2002     -0.000023  0.127208  0.00  1.000
## Year2003      0.054069  0.106452  0.51  0.612
## Year2004     -0.117616  0.139929 -0.84  0.401
## Year2005     -0.087129  0.099610 -0.87  0.383
## Year2006     -0.154640  0.112496 -1.37  0.171
## Year2007      0.002821  0.100723  0.03  0.978
## Year2008     -0.140550  0.115469 -1.22  0.225
## Year2009      0.029362  0.105329  0.28  0.781
## Year2010     -0.142436  0.125132 -1.14  0.256
## Year2011     -0.132108  0.134125 -0.98  0.326
## Year2012     -0.191219  0.101226 -1.89  0.060 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.271
## Multiple R-squared:  0.218, Adjusted R-squared:  0.148
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 16 weights are ~= 1. The remaining 250 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0901 0.8820 0.9540 0.8980 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      3.76e-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 1.492 1 1.221
## LastAuthorFemale 1.695 1 1.302
## Year 2.424 16 1.028

```

Residuals from first and last author



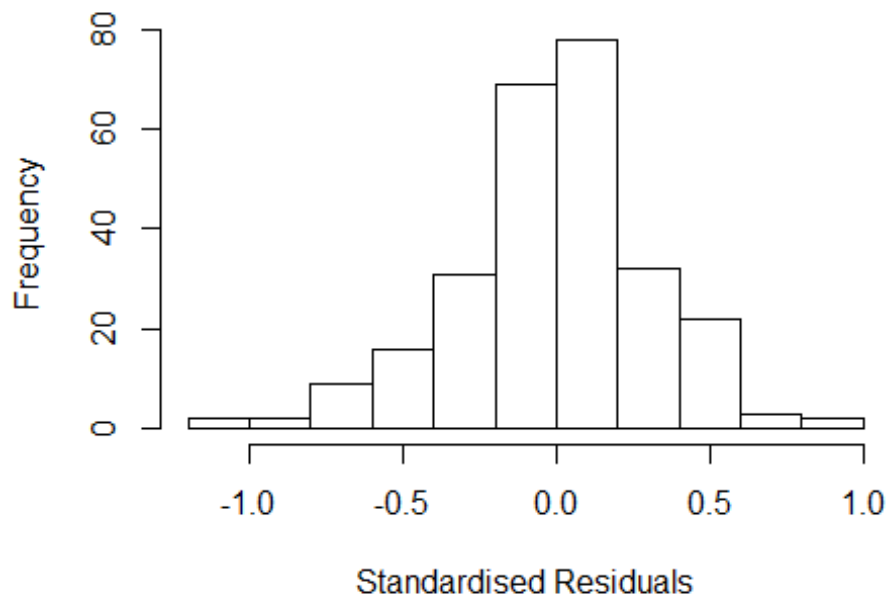
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.02064 -0.18587  0.00766  0.17269  0.85992
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.11243    0.08949   12.43  <2e-16 ***
## FirstAuthorFemale1  0.06877    0.04280    1.61   0.109
## LastAuthorFemale1  0.01368    0.05908    0.23   0.817
## Year1997          -0.01760    0.14157   -0.12   0.901
## Year1998          -0.10547    0.14073   -0.75   0.454
## Year1999          -0.07299    0.10642   -0.69   0.493
## Year2000          -0.11881    0.14302   -0.83   0.407
## Year2001          -0.03235    0.13124   -0.25   0.805
## Year2002          -0.12518    0.13536   -0.92   0.356
## Year2003          -0.02441    0.10180   -0.24   0.811
## Year2004          -0.17414    0.14878   -1.17   0.243
## Year2005          -0.06502    0.11009   -0.59   0.555
```

```

## Year2006      -0.22856    0.13097   -1.75    0.082 .
## Year2007      0.00739    0.10784    0.07    0.945
## Year2008     -0.15876    0.12475   -1.27    0.204
## Year2009     -0.03427    0.10467   -0.33    0.744
## Year2010     -0.20351    0.13301   -1.53    0.127
## Year2011     -0.18805    0.12640   -1.49    0.138
## Year2012     -0.20829    0.10604   -1.96    0.051 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.286
## Multiple R-squared:  0.0716, Adjusted R-squared:  0.00393
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 240 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.177  0.838   0.956   0.888   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      3.76e-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 1.45 1      1.204
## Year              1.45 16      1.012

```

Residuals from first author



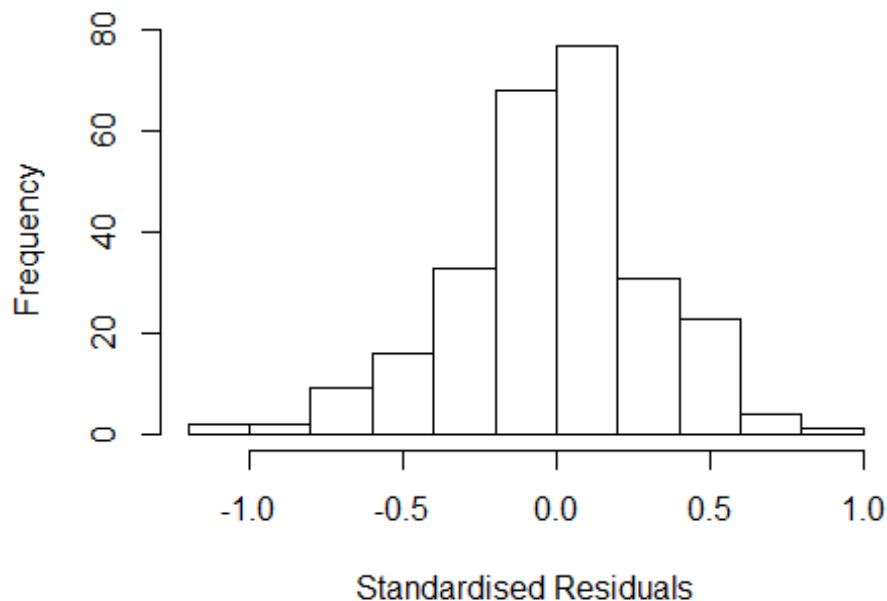
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0057 -0.1892 0.0101 0.1732 0.8588
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1122 0.0894 12.44 <2e-16 ***
## FirstAuthorFemale1 0.0701 0.0423 1.66 0.099 .
## Year1997 -0.0148 0.1405 -0.11 0.916
## Year1998 -0.1065 0.1425 -0.75 0.456
## Year1999 -0.0727 0.1063 -0.68 0.494
## Year2000 -0.1186 0.1426 -0.83 0.407
## Year2001 -0.0310 0.1309 -0.24 0.813
## Year2002 -0.1227 0.1353 -0.91 0.365
## Year2003 -0.0210 0.1015 -0.21 0.837
## Year2004 -0.1745 0.1481 -1.18 0.240
## Year2005 -0.0628 0.1103 -0.57 0.570
## Year2006 -0.2270 0.1314 -1.73 0.085 .
```

```

## Year2007          0.0111      0.1065      0.10      0.917
## Year2008          -0.1526     0.1188     -1.28     0.200
## Year2009          -0.0320     0.1038     -0.31     0.758
## Year2010          -0.1976     0.1310     -1.51     0.133
## Year2011          -0.1854     0.1261     -1.47     0.143
## Year2012          -0.2047     0.1044     -1.96     0.051 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.289
## Multiple R-squared:  0.0711, Adjusted R-squared:  0.00746
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 241 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.200  0.846  0.957  0.890  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      3.76e-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 1.659 1      1.288
## Year              1.659 16      1.016

```

Residuals from last author



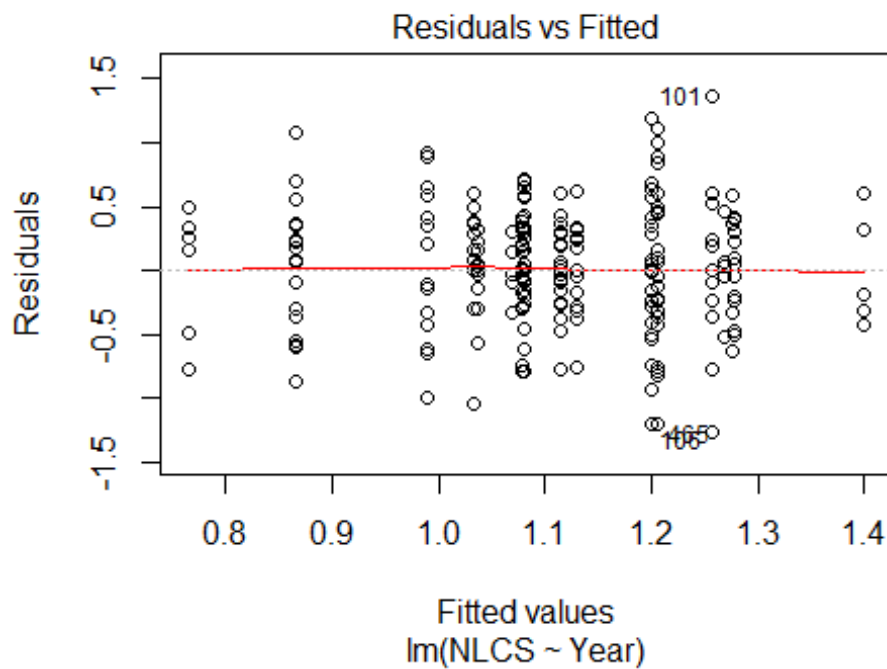
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.02945 -0.18221  0.00587  0.17210  0.84979
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.11249    0.08953   12.43  <2e-16 ***
## LastAuthorFemale1 0.02269    0.05752    0.39   0.694
## Year1997         0.00582    0.14354    0.04   0.968
## Year1998        -0.10573    0.14061   -0.75   0.453
## Year1999        -0.05106    0.10452   -0.49   0.626
## Year2000        -0.10714    0.14316   -0.75   0.455
## Year2001        -0.02228    0.13207   -0.17   0.866
## Year2002        -0.09335    0.12965   -0.72   0.472
## Year2003        -0.01154    0.10184   -0.11   0.910
## Year2004        -0.13545    0.14792   -0.92   0.361
## Year2005        -0.03616    0.10640   -0.34   0.734
## Year2006        -0.21242    0.12860   -1.65   0.100 .
```

```

## Year2007      0.02785      0.10760      0.26      0.796
## Year2008     -0.13155      0.12502     -1.05      0.294
## Year2009     -0.00491      0.10410     -0.05      0.962
## Year2010     -0.18034      0.12860     -1.40      0.162
## Year2011     -0.16518      0.12813     -1.29      0.199
## Year2012     -0.18459      0.10649     -1.73      0.084 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.286
## Multiple R-squared:  0.0611, Adjusted R-squared:  -0.00325
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 235 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.166  0.834  0.952  0.884  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      3.76e-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: 266"
## [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
##   12   11   14   23   18   14   21   13   15   24   25   24   23   33   21
## 2011 2012
##   33   37
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    5    5   12   12    6   11    5   11   19   19   16   19   23   13
## 2011 2012

```

```
## 25 27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 4 4 4 11 11 6 11 5 10 17 19 13 17 20 13
## 2011 2012
## 21 21
## [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
```

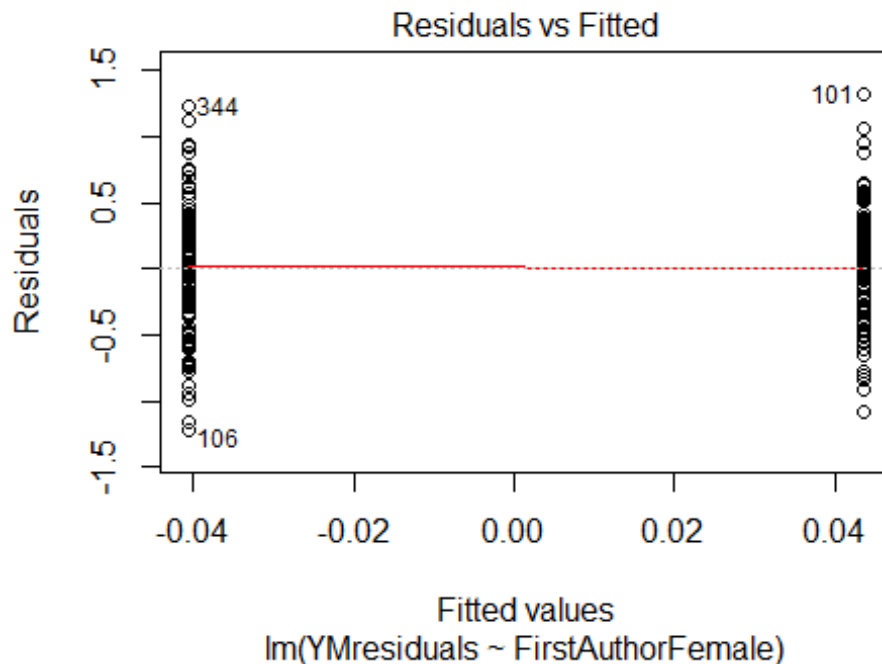


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.19, df = 1, p-value = 0.7
## [1] "Female first author team size 2018 geometric mean: 4.10641288039592"
## [1] "Male first author team size 2018 geometric mean: 3.47958137229683"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



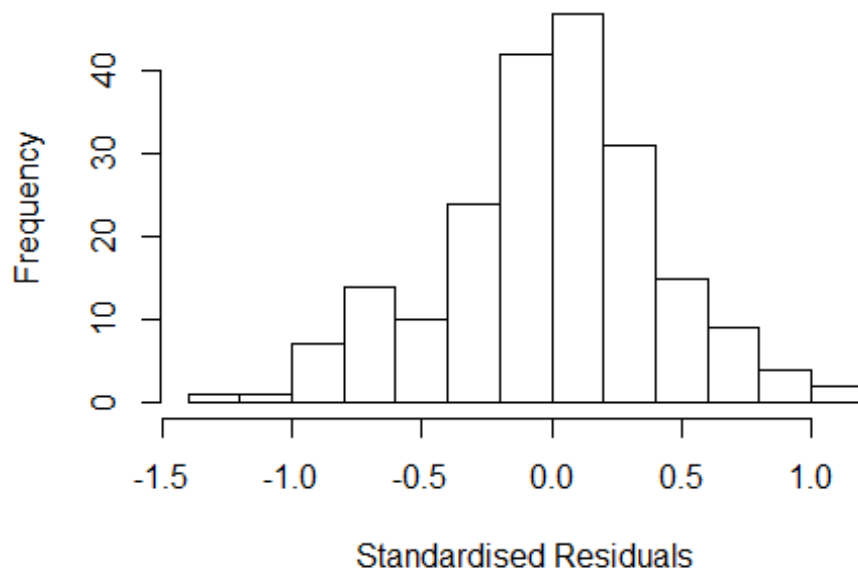
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 100, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.52274626078101"
## [1] "Male last author team size 2018 geometric mean: 3.89361716536627"

## 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 = 77, 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  2.183  1      1.477
## LastAuthorFemale  2.449  1      1.565
## UniqueAuthors    10.881  4      1.348
## Year              31.632 16      1.114
```

Residuals from first and last author and team size



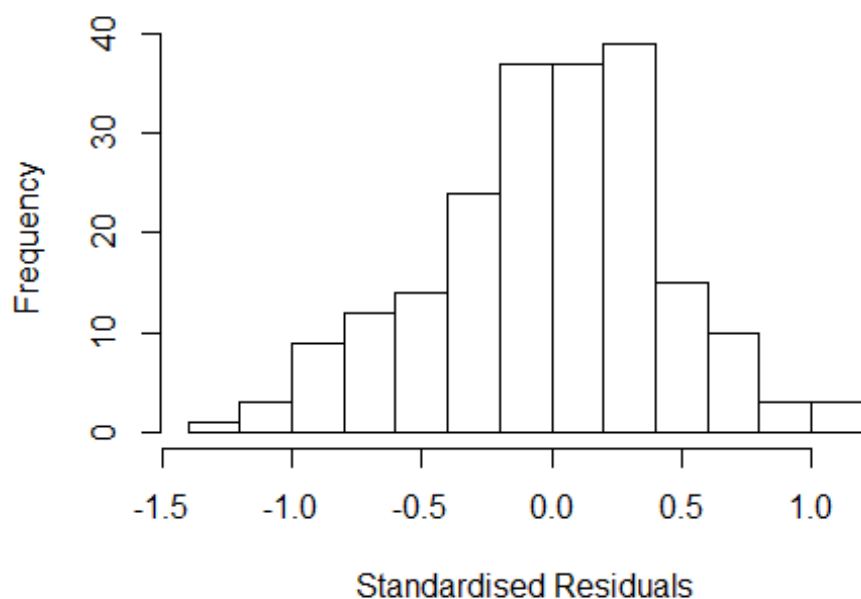
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3938 -0.2478  0.0173  0.2656  1.1102
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1165    0.2215   5.04 1.1e-06 ***
## FirstAuthorFemale1 -0.0445    0.0752  -0.59 0.55436
## LastAuthorFemale1  0.0577    0.0837   0.69 0.49170
## UniqueAuthors2     0.1197    0.1797   0.67 0.50644
## UniqueAuthors3     0.4730    0.1711   2.76 0.00629 **
## UniqueAuthors4     0.5129    0.1620   3.17 0.00181 **
## UniqueAuthors5     0.5722    0.1705   3.36 0.00096 ***
## Year1997          -0.2553    0.2297  -1.11 0.26799
## Year1998          -0.3951    0.2722  -1.45 0.14831
## Year1999          -0.1148    0.2407  -0.48 0.63396
```

```

## Year2000          -0.3504      0.2579    -1.36   0.17598
## Year2001          -0.4896      0.3509    -1.40   0.16460
## Year2002          -0.3251      0.3018    -1.08   0.28282
## Year2003          -0.3405      0.2326    -1.46   0.14491
## Year2004          -0.5477      0.2425    -2.26   0.02509 *
## Year2005          -0.4753      0.2302    -2.07   0.04032 *
## Year2006          -0.6220      0.2339    -2.66   0.00853 **
## Year2007          -0.5600      0.2802    -2.00   0.04711 *
## Year2008          -0.4982      0.2495    -2.00   0.04735 *
## Year2009          -0.3080      0.2561    -1.20   0.23064
## Year2010          -0.5175      0.2444    -2.12   0.03561 *
## Year2011          -0.3845      0.2632    -1.46   0.14582
## Year2012          -0.5988      0.2559    -2.34   0.02033 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.412
## Multiple R-squared:  0.186, Adjusted R-squared:  0.0884
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 190 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.230  0.857  0.956  0.901  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      4.83e-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 1.779 1      1.334
## LastAuthorFemale  1.835 1      1.355
## Year              2.431 16      1.028

```

Residuals from first and last author



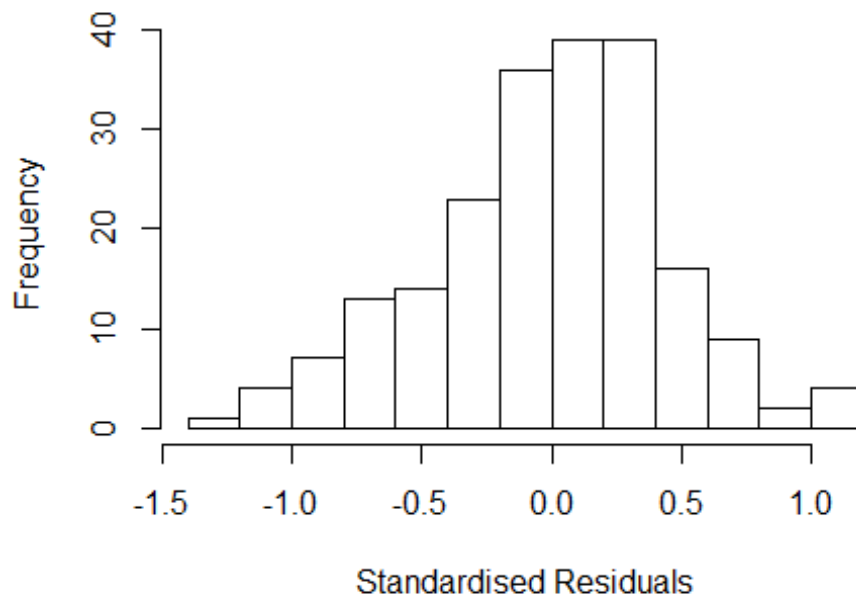
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2212 -0.2931  0.0309  0.2819  1.1339
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3111     0.3197   4.10 6.1e-05 ***
## FirstAuthorFemale1 -0.0428     0.0776  -0.55   0.58
## LastAuthorFemale1  -0.0319     0.0829  -0.38   0.70
## Year1997          -0.1310     0.3406  -0.38   0.70
## Year1998          -0.0250     0.3501  -0.07   0.94
## Year1999           0.0261     0.3298   0.08   0.94
## Year2000          -0.1180     0.3561  -0.33   0.74
## Year2001          -0.4871     0.3807  -1.28   0.20
## Year2002          -0.1488     0.3548  -0.42   0.68
## Year2003          -0.1730     0.3376  -0.51   0.61
## Year2004          -0.2691     0.3285  -0.82   0.41
## Year2005          -0.1701     0.3277  -0.52   0.60
```

```

## Year2006          -0.4229      0.3407    -1.24      0.22
## Year2007          -0.2418      0.3701    -0.65      0.51
## Year2008          -0.1138      0.3312    -0.34      0.73
## Year2009          -0.0151      0.3442    -0.04      0.96
## Year2010          -0.1386      0.3342    -0.41      0.68
## Year2011          -0.0989      0.3506    -0.28      0.78
## Year2012          -0.2114      0.3374    -0.63      0.53
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.0773, Adjusted R-squared:  -0.0111
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 185 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.433  0.865   0.953   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      4.83e-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 1.659 1      1.288
## Year              1.659 16      1.016

```

Residuals from first author



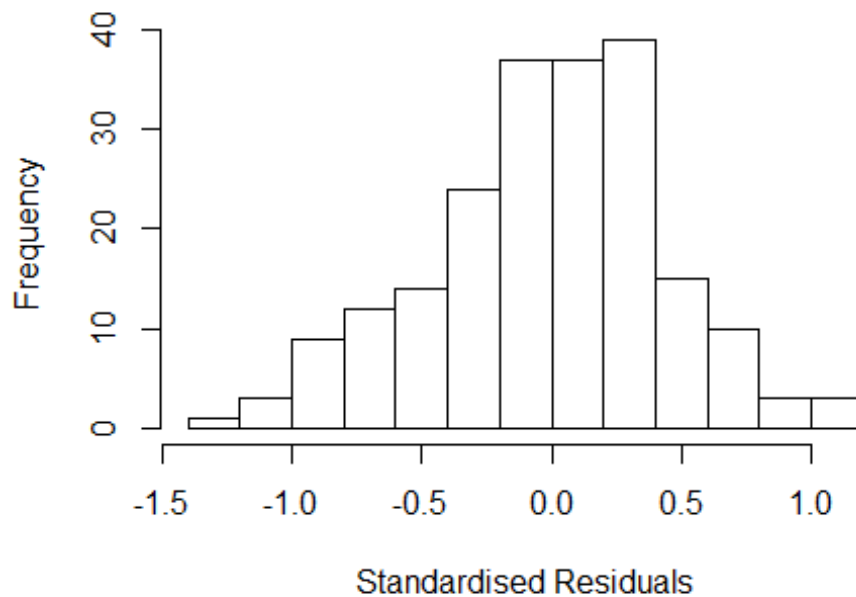
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2389 -0.2947  0.0207  0.2873  1.1481
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3055     0.3123   4.18 4.4e-05 ***
## FirstAuthorFemale1 -0.0503     0.0747  -0.67   0.50
## Year1997          -0.1317     0.3331  -0.40   0.69
## Year1998          -0.0408     0.3435  -0.12   0.91
## Year1999           0.0301     0.3226   0.09   0.93
## Year2000          -0.1129     0.3499  -0.32   0.75
## Year2001          -0.4887     0.3756  -1.30   0.19
## Year2002          -0.1423     0.3466  -0.41   0.68
## Year2003          -0.1851     0.3297  -0.56   0.58
## Year2004          -0.2638     0.3210  -0.82   0.41
## Year2005          -0.1667     0.3204  -0.52   0.60
## Year2006          -0.4259     0.3342  -1.27   0.20
```

```

## Year2007          -0.2472      0.3642   -0.68      0.50
## Year2008          -0.1195      0.3236   -0.37      0.71
## Year2009          -0.0164      0.3376   -0.05      0.96
## Year2010          -0.1460      0.3270   -0.45      0.66
## Year2011          -0.1052      0.3438   -0.31      0.76
## Year2012          -0.2134      0.3310   -0.64      0.52
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.0772, Adjusted R-squared:  -0.00584
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 184 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.416  0.863  0.954  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      4.83e-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 1.692 1      1.301
## Year              1.692 16      1.017

```

Residuals from last author



```
## [1] "List of 0 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
## -1.2328 -0.3024 0.0136 0.2805 1.1628
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2942 0.3257 3.97 0.0001 ***
## LastAuthorFemale1 -0.0425 0.0800 -0.53 0.5959
## Year1997 -0.1228 0.3465 -0.35 0.7235
## Year1998 -0.0231 0.3561 -0.07 0.9482
## Year1999 0.0362 0.3359 0.11 0.9143
## Year2000 -0.1083 0.3608 -0.30 0.7643
## Year2001 -0.4795 0.3865 -1.24 0.2163
## Year2002 -0.1574 0.3591 -0.44 0.6617
## Year2003 -0.1905 0.3411 -0.56 0.5772
## Year2004 -0.2692 0.3341 -0.81 0.4214
## Year2005 -0.1708 0.3342 -0.51 0.6099
## Year2006 -0.4226 0.3472 -1.22 0.2250
```

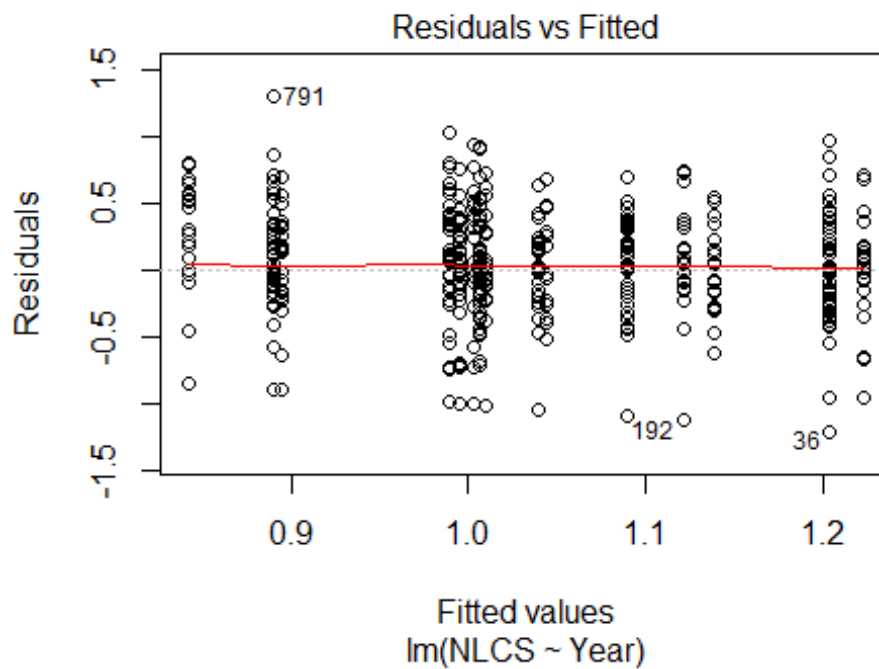


```

## Year2007          -0.2484      0.3767   -0.66   0.5105
## Year2008          -0.1222      0.3354   -0.36   0.7159
## Year2009          -0.0188      0.3506   -0.05   0.9573
## Year2010          -0.1352      0.3405   -0.40   0.6917
## Year2011          -0.1005      0.3565   -0.28   0.7784
## Year2012          -0.2159      0.3414   -0.63   0.5278
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.0753, Adjusted R-squared:  -0.00787
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 185 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.417  0.858  0.953  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      4.83e-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: 207"
## [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
##   25   36   32   63   35   53   46   32   36   39   32   31   32   36   47
## 2011 2012
##   69   73
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   17   22   21   24   23   19   29   22   27   35   24   19   23   25   29
## 2011 2012

```

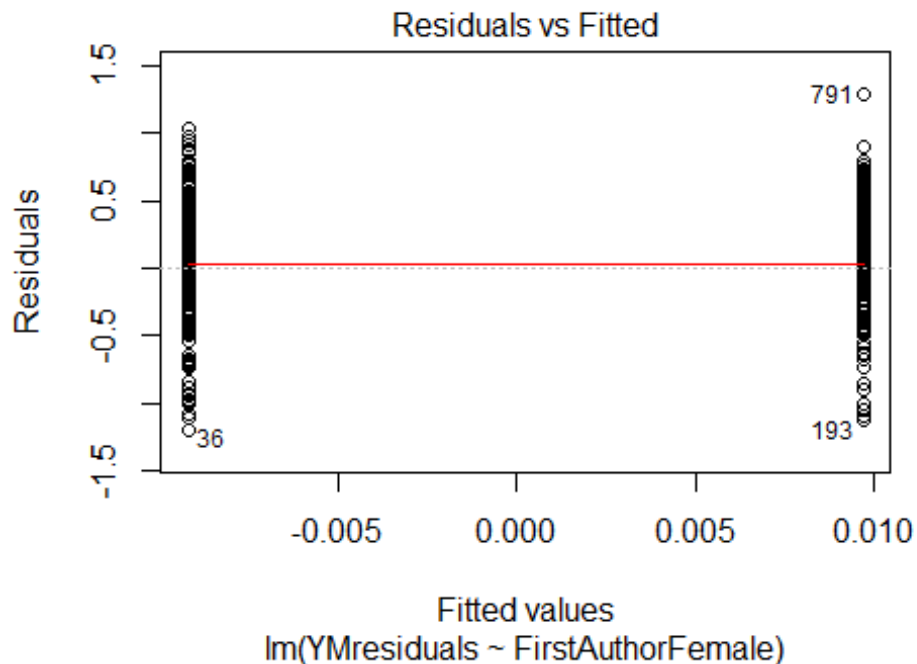
```
## 50 47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 16 19 19 22 21 18 29 20 24 34 23 18 21 25 27
## 2011 2012
## 42 39
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 28, df = 16, p-value = 0.03
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.0091, df = 1, p-value = 0.9
## [1] "Female first author team size 2018 geometric mean: 3.21732002714139"
## [1] "Male first author team size 2018 geometric mean: 3.52697788429511"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

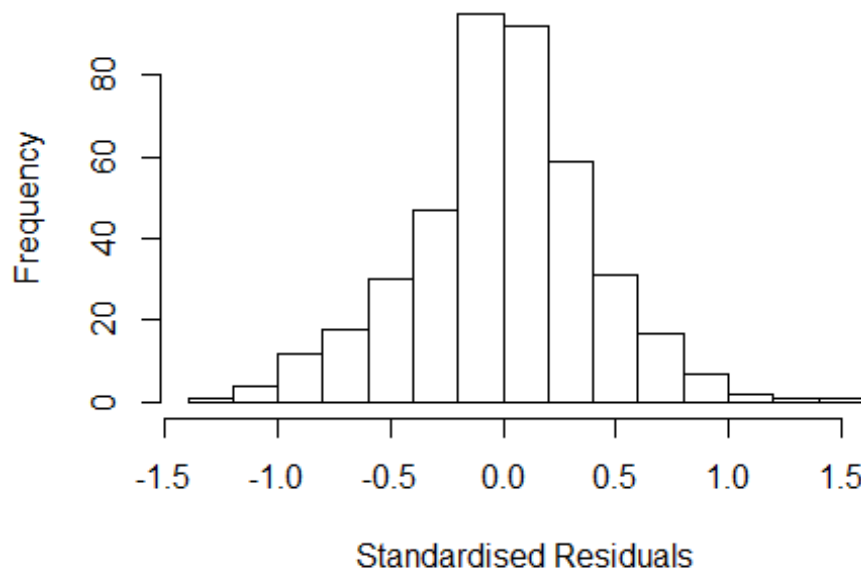
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 160, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.44968137869122"
## [1] "Male last author team size 2018 geometric mean: 3.2786215669812"

## 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 = 190, 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 1.195  1      1.093
## LastAuthorFemale  1.685  1      1.298
## UniqueAuthors    3.513  4      1.170
## Year              4.754 16      1.050
```

Residuals from first and last author and team size



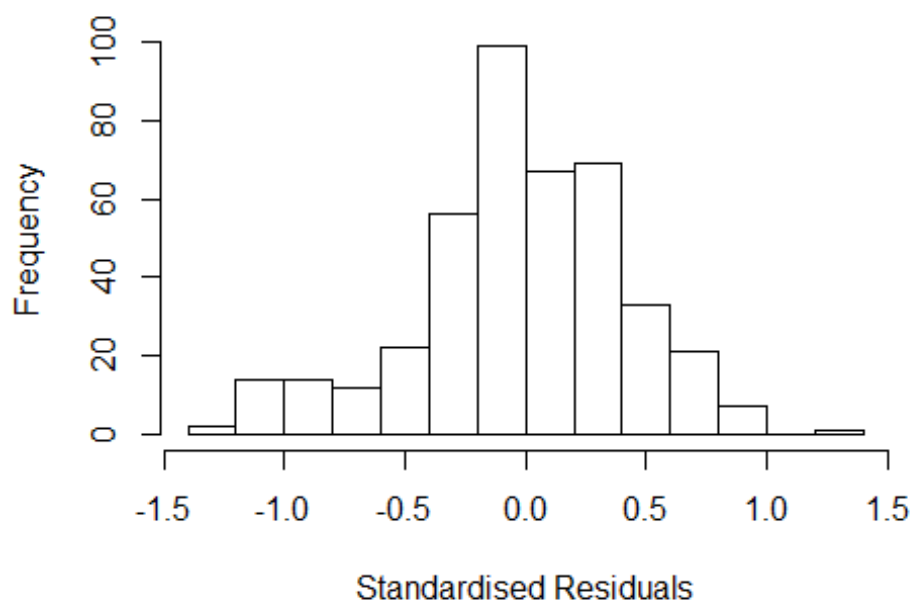
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21217 -0.23771  0.00673  0.24890  1.57799
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.70627    0.13627   5.18 3.5e-07 ***
## FirstAuthorFemale1 0.05216    0.04133   1.26 0.20767
## LastAuthorFemale1 -0.04260    0.05185  -0.82 0.41185
## UniqueAuthors2    0.29645    0.09715   3.05 0.00243 **
## UniqueAuthors3    0.31028    0.09140   3.39 0.00076 ***
## UniqueAuthors4    0.42800    0.08471   5.05 6.7e-07 ***
## UniqueAuthors5    0.48459    0.08076   6.00 4.5e-09 ***
## Year1997          0.21224    0.16737   1.27 0.20553
## Year1998          0.02966    0.14713   0.20 0.84035
## Year1999          0.20585    0.13891   1.48 0.13916
```

```

## Year2000      0.12699    0.15513    0.82  0.41351
## Year2001      0.01433    0.19739    0.07  0.94215
## Year2002     -0.14417    0.16920   -0.85  0.39469
## Year2003      0.07114    0.12243    0.58  0.56151
## Year2004     -0.03085    0.13242   -0.23  0.81590
## Year2005      0.12218    0.12131    1.01  0.31445
## Year2006     -0.04830    0.14251   -0.34  0.73481
## Year2007      0.02504    0.13578    0.18  0.85379
## Year2008     -0.07457    0.15697   -0.48  0.63504
## Year2009     -0.00993    0.12041   -0.08  0.93429
## Year2010     -0.05559    0.12466   -0.45  0.65593
## Year2011     -0.10482    0.13927   -0.75  0.45209
## Year2012     -0.08948    0.11982   -0.75  0.45564
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.36
## Multiple R-squared:  0.224, Adjusted R-squared:  0.181
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 40 weights are ~= 1. The remaining 377 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0155 0.8460 0.9510 0.8820 0.9850 0.9990
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          2.40e-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 1.135 1 1.065
## LastAuthorFemale 1.594 1 1.263
## Year 1.683 16 1.016

```

Residuals from first and last author



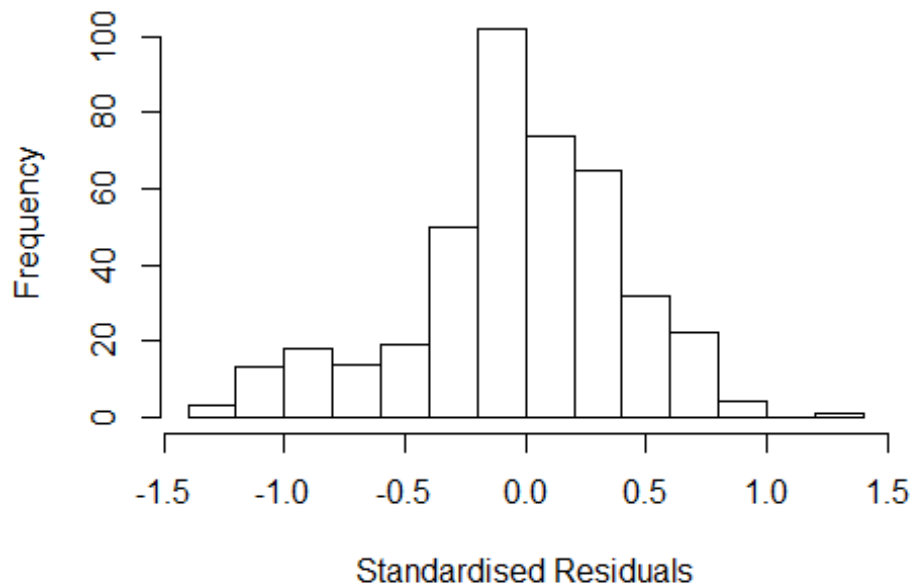
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2398 -0.2381 -0.0158 0.2617 1.3476
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0181 0.1069 9.53 <2e-16 ***
## FirstAuthorFemale1 0.0842 0.0430 1.96 0.051 .
## LastAuthorFemale1 -0.1181 0.0560 -2.11 0.035 *
## Year1997 0.2217 0.1493 1.49 0.138
## Year1998 0.0510 0.1433 0.36 0.722
## Year1999 0.2521 0.1291 1.95 0.052 .
## Year2000 0.1980 0.1585 1.25 0.213
## Year2001 0.0755 0.2200 0.34 0.732
## Year2002 -0.1257 0.1861 -0.68 0.500
## Year2003 0.1068 0.1272 0.84 0.401
## Year2004 -0.1334 0.1312 -1.02 0.310
## Year2005 0.1083 0.1210 0.90 0.371
```

```

## Year2006          0.0537      0.1493      0.36      0.719
## Year2007          0.1512      0.1299      1.16      0.245
## Year2008         -0.0248      0.1474     -0.17      0.867
## Year2009          0.0460      0.1186      0.39      0.699
## Year2010          0.0287      0.1240      0.23      0.817
## Year2011         -0.1427      0.1308     -1.09      0.276
## Year2012         -0.0828      0.1156     -0.72      0.474
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.384
## Multiple R-squared:  0.0938, Adjusted R-squared:  0.0528
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 389 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.192  0.855   0.957   0.885   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      2.40e-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 1.132 1      1.064
## Year              1.132 16      1.004

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2758 -0.2361 -0.0163  0.2504  1.2691
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9940    0.1046   9.50  <2e-16 ***
## FirstAuthorFemale1 0.0522    0.0439   1.19   0.235
## Year1997        0.1962    0.1456   1.35   0.178
## Year1998        0.0421    0.1405   0.30   0.765
## Year1999        0.2773    0.1244   2.23   0.026 *
## Year2000        0.2296    0.1557   1.47   0.141
## Year2001        0.0927    0.2307   0.40   0.688
## Year2002       -0.1009    0.1842  -0.55   0.584
## Year2003        0.1080    0.1274   0.85   0.397
## Year2004       -0.1538    0.1333  -1.15   0.249
## Year2005        0.1181    0.1198   0.99   0.325
## Year2006        0.0614    0.1524   0.40   0.687
```

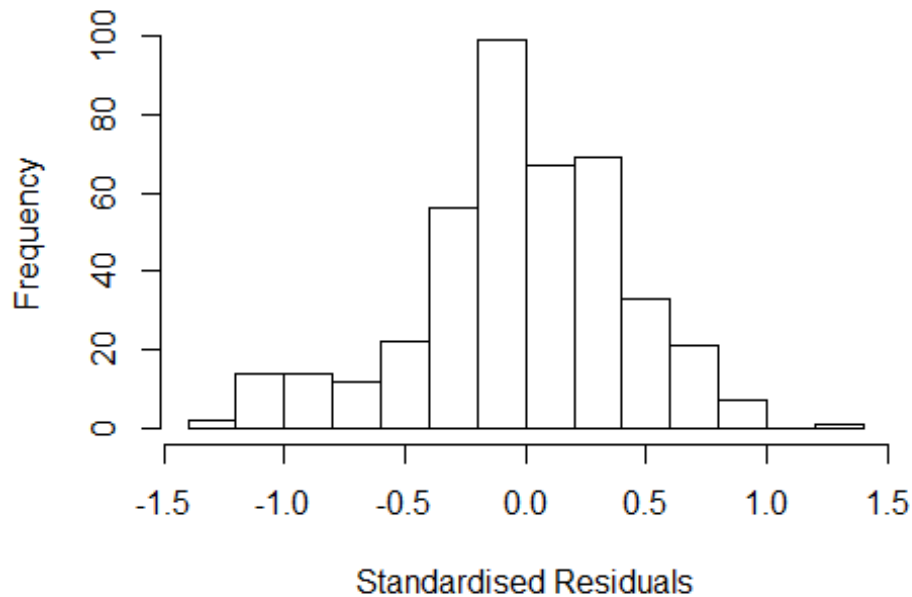


```

## Year2007          0.1390      0.1332      1.04      0.297
## Year2008         -0.0194      0.1493     -0.13      0.897
## Year2009          0.0357      0.1163      0.31      0.759
## Year2010          0.0083      0.1216      0.07      0.946
## Year2011         -0.1262      0.1309     -0.96      0.335
## Year2012         -0.0846      0.1141     -0.74      0.459
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.383
## Multiple R-squared:  0.0795, Adjusted R-squared:  0.0402
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 379 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.245  0.845   0.946   0.880   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      2.40e-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 1.527 1      1.236
## Year              1.527 16      1.013

```

Residuals from last author



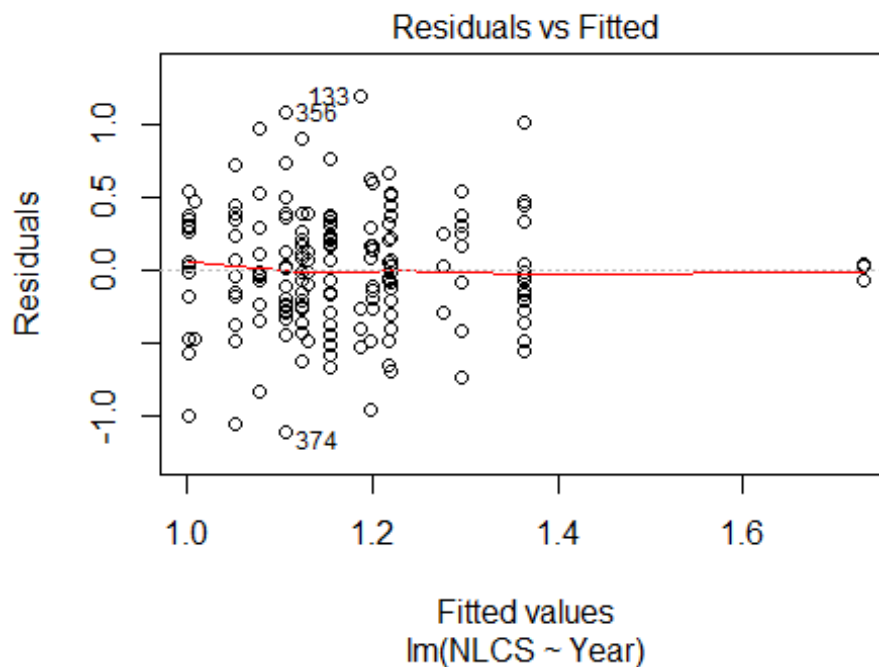
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2476 -0.2504 -0.0249 0.2656 1.3613
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0399 0.0995 10.45 <2e-16 ***
## LastAuthorFemale1 -0.0906 0.0549 -1.65 0.100 .
## Year1997 0.2078 0.1447 1.44 0.152
## Year1998 0.0527 0.1422 0.37 0.711
## Year1999 0.2531 0.1221 2.07 0.039 *
## Year2000 0.1966 0.1526 1.29 0.198
## Year2001 0.0694 0.2096 0.33 0.741
## Year2002 -0.1267 0.1798 -0.70 0.481
## Year2003 0.1240 0.1237 1.00 0.317
## Year2004 -0.1093 0.1269 -0.86 0.390
## Year2005 0.1276 0.1153 1.11 0.269
## Year2006 0.0663 0.1469 0.45 0.652
```

```

## Year2007          0.1563      0.1275      1.23      0.221
## Year2008          -0.0123     0.1472     -0.08     0.933
## Year2009          0.0608     0.1140     0.53     0.594
## Year2010          0.0335     0.1201     0.28     0.781
## Year2011          -0.1216     0.1262     -0.96     0.336
## Year2012          -0.0649     0.1105     -0.59     0.557
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.0849, Adjusted R-squared:  0.0459
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 385 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.207  0.856  0.954  0.888  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      2.40e-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: 417"
## [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
##    9   20   35   30   10   24   14   21   11   22   15   17   23   11   22
## 2011 2012
##   23   24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    3    8   15    2    4    7   13    9   14    8   12   17    6   17
## 2011 2012

```

```
## 17 20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 3 3 7 13 2 4 5 12 8 9 5 12 14 6 17
## 2011 2012
## 13 17
## [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
```



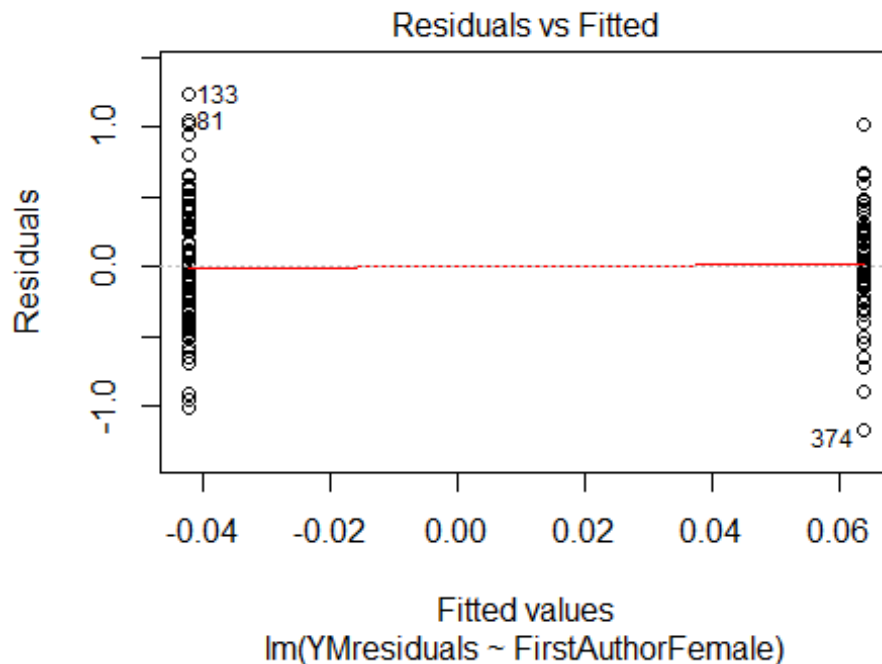
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.69, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 4.07778280532755"
## [1] "Male first author team size 2018 geometric mean: 1.73205080756888"

## 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 = 20, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.3345372167932"
## [1] "Male last author team size 2018 geometric mean: 2.78315768371374"

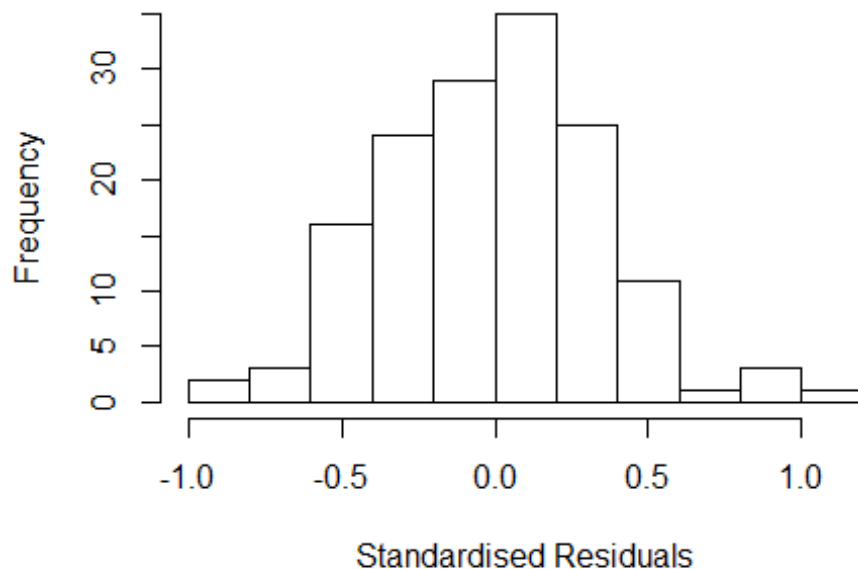
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 36, 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	1.949	1	1.396
## LastAuthorFemale	1.922	1	1.386
## UniqueAuthors	9.833	4	1.331
## Year	19.603	16	1.097

Residuals from first and last author and team size



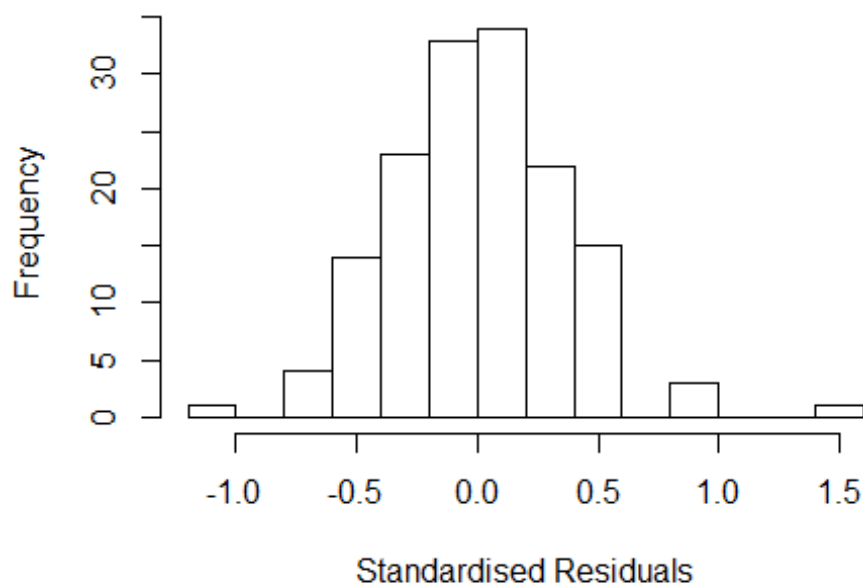
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9069 -0.2268 0.0103 0.2418 1.1436
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08479 0.20401 5.32 4.6e-07 ***
## FirstAuthorFemale1 0.06962 0.06613 1.05 0.2944
## LastAuthorFemale1 0.19854 0.07193 2.76 0.0066 **
## UniqueAuthors2 0.03972 0.18705 0.21 0.8322
## UniqueAuthors3 -0.15899 0.19441 -0.82 0.4150
## UniqueAuthors4 -0.17239 0.20004 -0.86 0.3904
## UniqueAuthors5 0.02461 0.19655 0.13 0.9005
## Year1997 0.59870 0.18122 3.30 0.0012 **
## Year1998 0.11000 0.16447 0.67 0.5048
## Year1999 0.16999 0.15237 1.12 0.2667
```

```

## Year2000      -0.11651    0.52685   -0.22    0.8253
## Year2001      -0.08766    0.34769   -0.25    0.8014
## Year2002       0.37686    0.22897    1.65    0.1023
## Year2003      -0.00548    0.15270   -0.04    0.9714
## Year2004       0.28114    0.21680    1.30    0.1971
## Year2005       0.02095    0.17553    0.12    0.9052
## Year2006       0.12914    0.17763    0.73    0.4686
## Year2007      -0.06658    0.19629   -0.34    0.7350
## Year2008      -0.02226    0.14840   -0.15    0.8810
## Year2009       0.02842    0.17945    0.16    0.8744
## Year2010       0.11338    0.14653    0.77    0.4405
## Year2011      -0.05497    0.16977   -0.32    0.7467
## Year2012      -0.01614    0.13842   -0.12    0.9073
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.348
## Multiple R-squared:  0.239, Adjusted R-squared:  0.107
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 136 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.258  0.874  0.953  0.907  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      6.67e-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 1.853 1      1.361
## LastAuthorFemale  1.706 1      1.306
## Year              2.477 16      1.029

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.018692 -0.216205 0.000204 0.240456 1.440640
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0614 0.1790 5.93 2.5e-08 ***
## FirstAuthorFemale1 0.0832 0.0666 1.25 0.2141
## LastAuthorFemale1 0.1915 0.0735 2.61 0.0102 *
## Year1997 0.5771 0.1849 3.12 0.0022 **
## Year1998 0.0889 0.2325 0.38 0.7029
## Year1999 0.1610 0.1990 0.81 0.4201
## Year2000 -0.0534 0.5542 -0.10 0.9234
## Year2001 -0.3145 0.2503 -1.26 0.2111
## Year2002 0.2310 0.2541 0.91 0.3650
## Year2003 -0.0427 0.2027 -0.21 0.8335
## Year2004 0.2871 0.2335 1.23 0.2211
## Year2005 -0.0527 0.2179 -0.24 0.8094
```

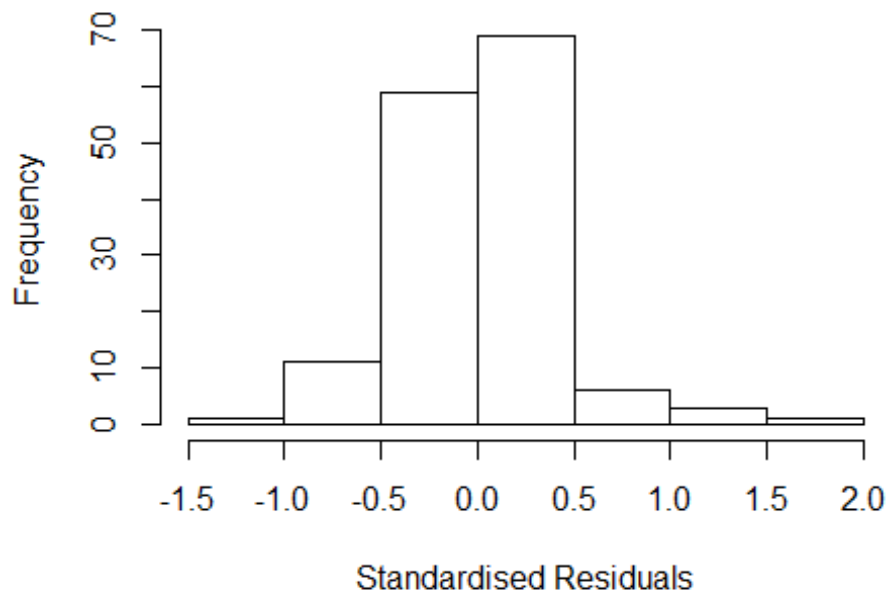


```

## Year2006          0.0566      0.2252      0.25      0.8020
## Year2007          -0.1299      0.2125     -0.61      0.5419
## Year2008          -0.0308      0.1916     -0.16      0.8725
## Year2009          -0.0276      0.2029     -0.14      0.8921
## Year2010           0.0542      0.1897      0.29      0.7755
## Year2011          -0.0540      0.2239     -0.24      0.8096
## Year2012          -0.0671      0.1917     -0.35      0.7268
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.2,   Adjusted R-squared:  0.0897
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 140 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0372 0.8720 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      6.67e-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 1.799 1      1.341
## Year              1.799 16      1.019

```

Residuals from first author



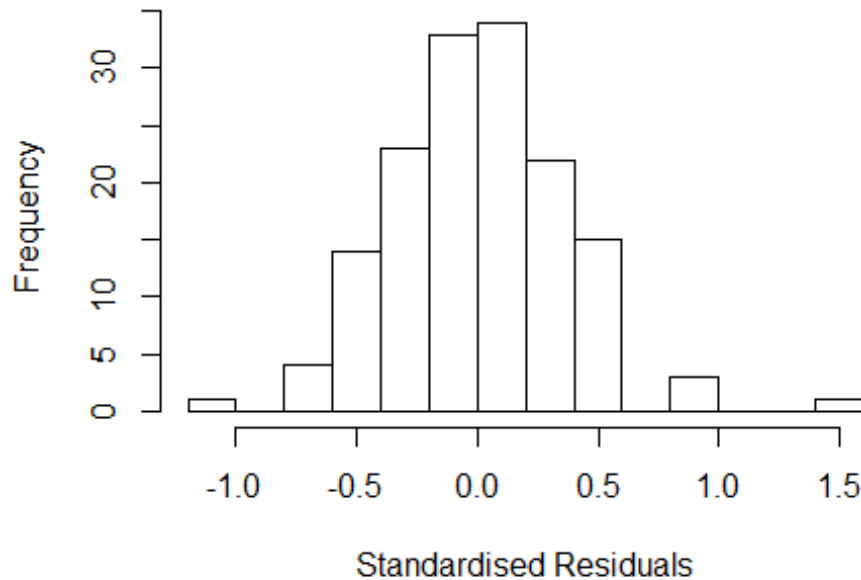
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0368 -0.2240 0.0249 0.2558 1.5904
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2425 0.1753 7.09 7.4e-11 ***
## FirstAuthorFemale1 0.1207 0.0655 1.84 0.067 .
## Year1997 0.4469 0.1786 2.50 0.014 *
## Year1998 -0.0538 0.2461 -0.22 0.827
## Year1999 0.0359 0.1955 0.18 0.854
## Year2000 -0.2345 0.5590 -0.42 0.676
## Year2001 -0.4539 0.1859 -2.44 0.016 *
## Year2002 0.0718 0.2545 0.28 0.778
## Year2003 -0.2057 0.2083 -0.99 0.325
## Year2004 0.1356 0.2342 0.58 0.563
## Year2005 -0.2121 0.2149 -0.99 0.326
## Year2006 -0.0801 0.2350 -0.34 0.734
```

```

## Year2007          -0.2531      0.2110    -1.20      0.233
## Year2008          -0.1662      0.1907    -0.87      0.385
## Year2009          -0.1618      0.2192    -0.74      0.462
## Year2010          -0.0631      0.1958    -0.32      0.748
## Year2011          -0.1900      0.2286    -0.83      0.407
## Year2012          -0.1869      0.2048    -0.91      0.363
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.338
## Multiple R-squared:  0.16,   Adjusted R-squared:  0.0514
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## observation 30 is an outlier with |weight| = 0 ( < 0.00067);
## 12 weights are ~= 1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.273  0.855   0.944   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      6.67e-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 1.621 1          1.273
## Year          1.621 16          1.015

```

Residuals from last author



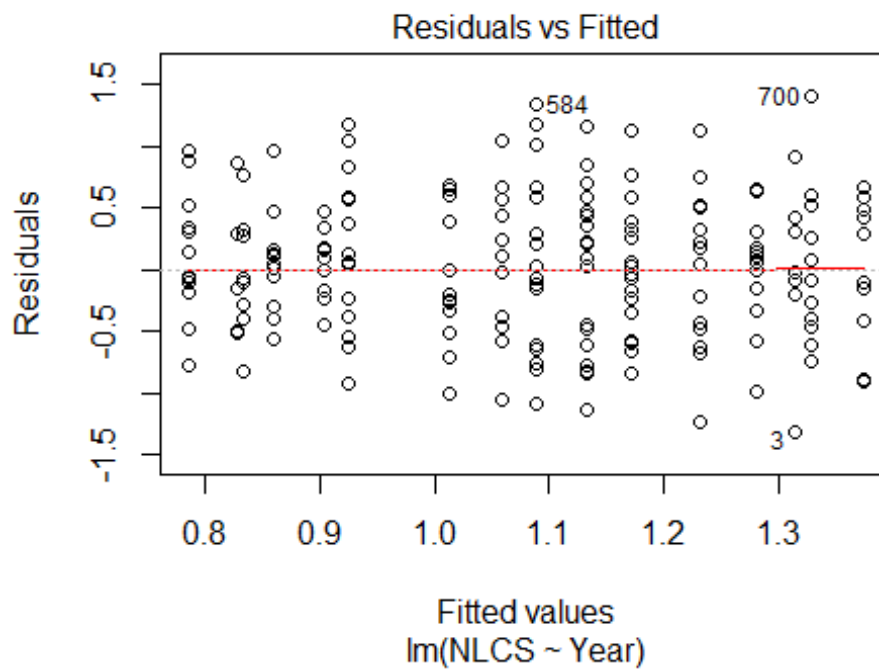
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.07417 -0.22028 -0.00195 0.24626 1.42477
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.068535 0.155248 6.88 2.1e-10 ***
## LastAuthorFemale1 0.208983 0.071038 2.94 0.00385 **
## Year1997 0.591487 0.153596 3.85 0.00018 ***
## Year1998 0.114344 0.222242 0.51 0.60776
## Year1999 0.169351 0.179276 0.94 0.34657
## Year2000 -0.060535 0.546469 -0.11 0.91196
## Year2001 -0.323287 0.239730 -1.35 0.17979
## Year2002 0.286372 0.219355 1.31 0.19399
## Year2003 0.005638 0.172608 0.03 0.97399
## Year2004 0.291802 0.212458 1.37 0.17194
## Year2005 -0.016337 0.193921 -0.08 0.93299
## Year2006 0.079299 0.202422 0.39 0.69587
```

```

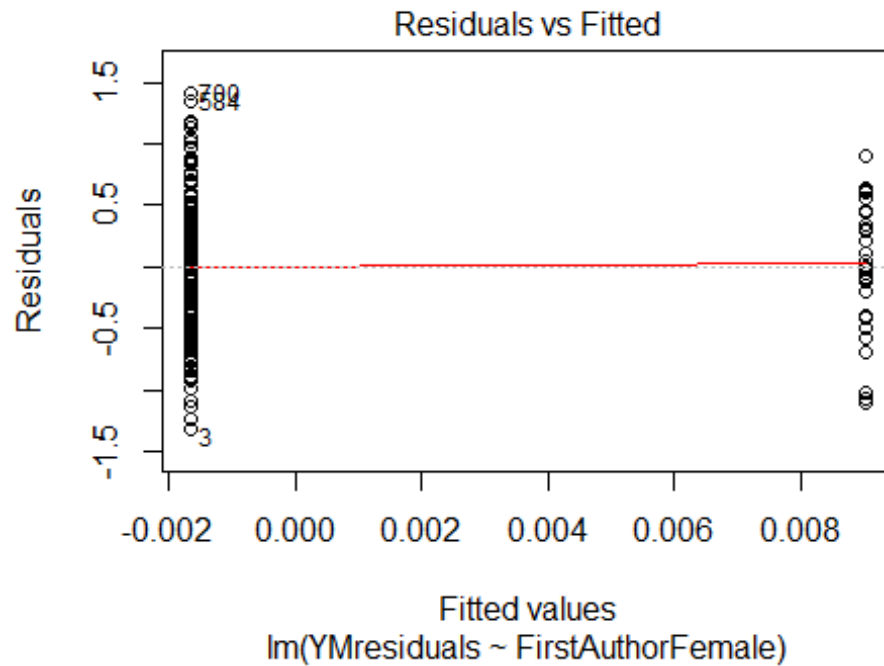
## Year2007          -0.108230    0.193885    -0.56    0.57764
## Year2008          -0.008902    0.165211    -0.05    0.95711
## Year2009          -0.000639    0.169563     0.00    0.99700
## Year2010           0.066874    0.164130     0.41    0.68434
## Year2011          -0.028000    0.207284    -0.14    0.89275
## Year2012          -0.037487    0.166363    -0.23    0.82207
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.342
## Multiple R-squared:  0.194, Adjusted R-squared:  0.0897
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0445 0.8630 0.9520 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      6.67e-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: 150"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##   18   26   28   24   30   27   27   44   30   48   30   32   32   27   35
## 2012
##   23
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    7   14   17    9    5   15   12   17   12   12   10   17   17   11   18
## 2012

```

```
## 12
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
## 7 14 16 9 5 12 9 17 10 12 8 15 16 11 16
## 2012
## 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 15, p-value = 0.4
```

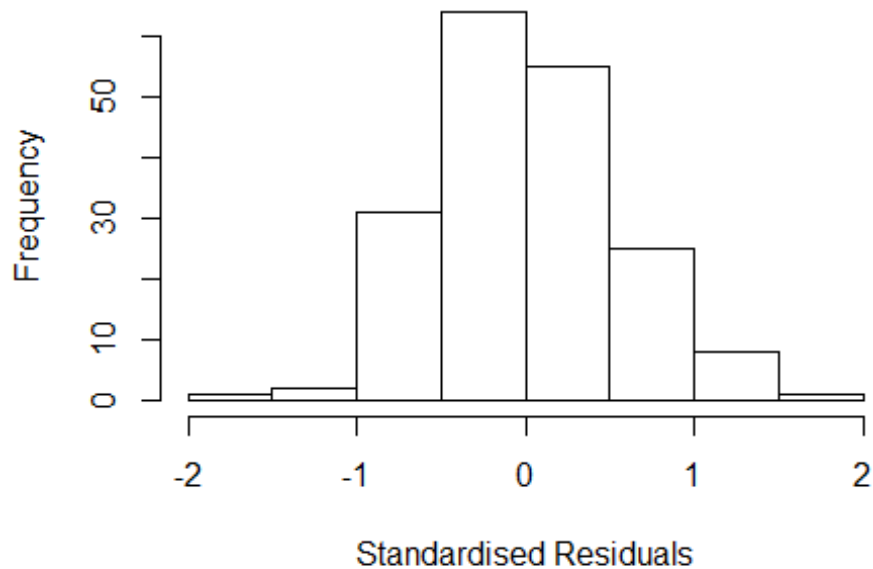


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.64, df = 1, p-value = 0.4
```



```
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 1.78038939137544"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.663 1          1.290
## LastAuthorFemale  1.518 1          1.232
## UniqueAuthors    3.940 4          1.187
## Year              6.131 15         1.062
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.5150 -0.4035 -0.0243  0.3983  1.6547
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1688    0.2324   5.03 1.3e-06 ***
## FirstAuthorFemale1 -0.0164    0.1156  -0.14  0.8876
## LastAuthorFemale1 -0.0302    0.1369  -0.22  0.8255
## UniqueAuthors2     0.3462    0.1317   2.63  0.0094 **
## UniqueAuthors3     0.1900    0.1448   1.31  0.1912
## UniqueAuthors4     0.2289    0.1909   1.20  0.2322
## UniqueAuthors5     0.0921    0.2790   0.33  0.7418
## Year1998          -0.2112    0.3009  -0.70  0.4838
## Year1999          -0.4252    0.2878  -1.48  0.1414
## Year2000          -0.4927    0.2741  -1.80  0.0741 .
```

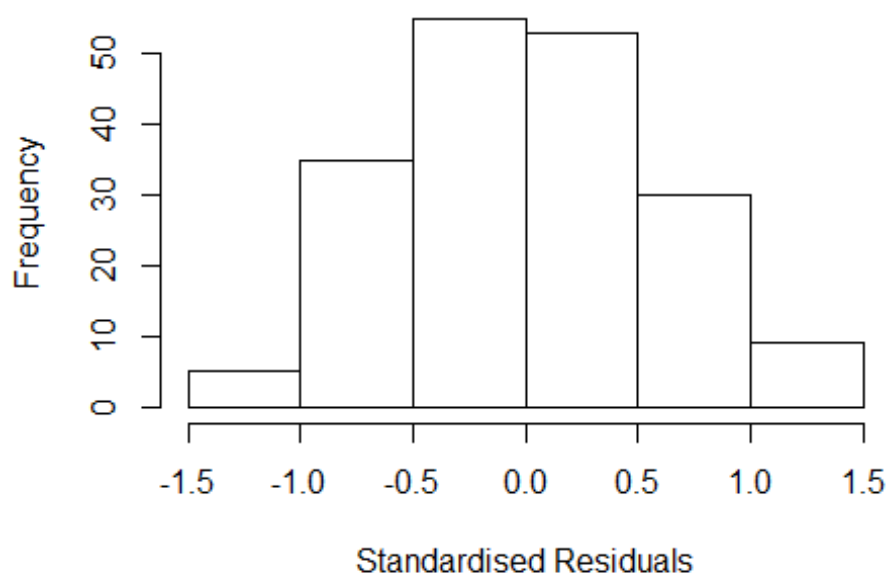


```

## Year2001          -0.5287      0.3279   -1.61    0.1088
## Year2002          -0.3361      0.2745   -1.22    0.2227
## Year2003          -0.2042      0.2704   -0.76    0.4511
## Year2004          -0.2414      0.2511   -0.96    0.3376
## Year2005          -0.0628      0.2801   -0.22    0.8229
## Year2006          -0.4856      0.2401   -2.02    0.0447 *
## Year2007           0.1673      0.3151    0.53    0.5962
## Year2008          -0.5796      0.2737   -2.12    0.0357 *
## Year2009          -0.2597      0.2959   -0.88    0.3814
## Year2010          -0.5151      0.2451   -2.10    0.0371 *
## Year2011          -0.3324      0.2838   -1.17    0.2433
## Year2012          -0.0885      0.3314   -0.27    0.7898
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.15,   Adjusted R-squared:  0.0424
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 170 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.442  0.896  0.956  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      5.35e-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 1.403 1      1.184
## LastAuthorFemale  1.409 1      1.187
## Year              1.968 15      1.023

```

Residuals from first and last author



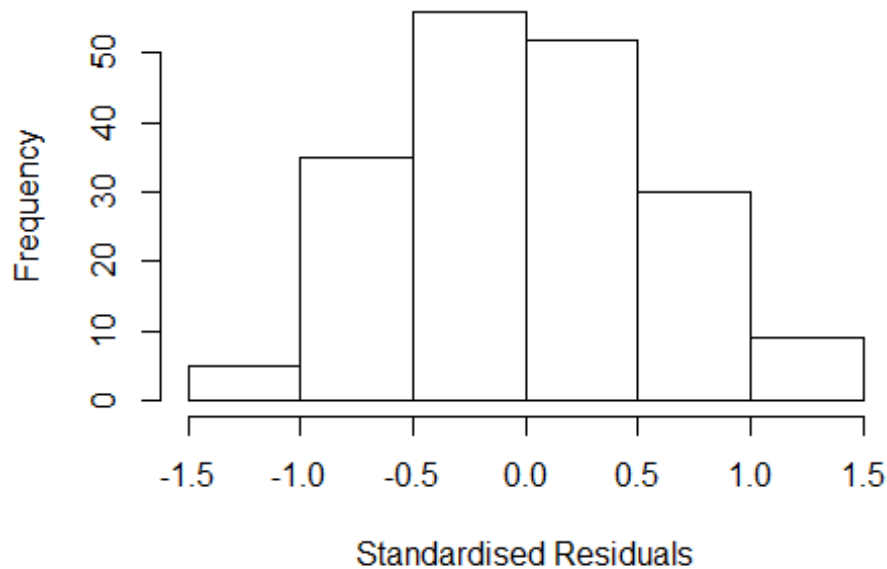
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3454 -0.4387 -0.0112 0.3860 1.4520
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3454 0.2248 5.98 1.3e-08 ***
## FirstAuthorFemale1 0.0754 0.1093 0.69 0.491
## LastAuthorFemale1 0.0547 0.1463 0.37 0.709
## Year1998 -0.1093 0.2886 -0.38 0.705
## Year1999 -0.5037 0.2930 -1.72 0.087 .
## Year2000 -0.5084 0.2733 -1.86 0.065 .
## Year2001 -0.5645 0.3507 -1.61 0.109
## Year2002 -0.3230 0.2756 -1.17 0.243
## Year2003 -0.1770 0.2909 -0.61 0.544
## Year2004 -0.2001 0.2599 -0.77 0.442
## Year2005 -0.0856 0.2786 -0.31 0.759
## Year2006 -0.5271 0.2556 -2.06 0.041 *
```

```

## Year2007          0.0650      0.3001      0.22      0.829
## Year2008         -0.5956      0.2832     -2.10      0.037 *
## Year2009         -0.2092      0.3074     -0.68      0.497
## Year2010         -0.4481      0.2395     -1.87      0.063 .
## Year2011         -0.2890      0.2919     -0.99      0.323
## Year2012         -0.0625      0.3082     -0.20      0.840
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.634
## Multiple R-squared:  0.103, Adjusted R-squared:  0.0124
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 172 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.580  0.894  0.953  0.924  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      5.35e-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 1.408 1          1.186
## Year              1.408 15          1.011

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3462 -0.4413 -0.0136 0.3857 1.4536
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3462 0.2242 6.00 1.1e-08 ***
## FirstAuthorFemale1 0.0771 0.1092 0.71 0.481
## Year1998 -0.1099 0.2884 -0.38 0.704
## Year1999 -0.5056 0.2932 -1.72 0.086 .
## Year2000 -0.5091 0.2728 -1.87 0.064 .
## Year2001 -0.5553 0.3455 -1.61 0.110
## Year2002 -0.3237 0.2753 -1.18 0.241
## Year2003 -0.1780 0.2906 -0.61 0.541
## Year2004 -0.1976 0.2586 -0.76 0.446
## Year2005 -0.0738 0.2695 -0.27 0.785
## Year2006 -0.5236 0.2542 -2.06 0.041 *
## Year2007 0.0638 0.3016 0.21 0.833
```

```

## Year2008          -0.5969      0.2832   -2.11    0.037 *
## Year2009          -0.2105      0.3074   -0.68    0.494
## Year2010          -0.4437      0.2383   -1.86    0.064 .
## Year2011          -0.2868      0.2929   -0.98    0.329
## Year2012          -0.0648      0.3081   -0.21    0.834
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.628
## Multiple R-squared:  0.103, Adjusted R-squared:  0.0184
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 176 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.571  0.893  0.954  0.924  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          5.35e-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 1.409 1          1.187
## Year              1.409 15          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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3654 -0.4406 -0.0112  0.3683  1.4404
##

```

```

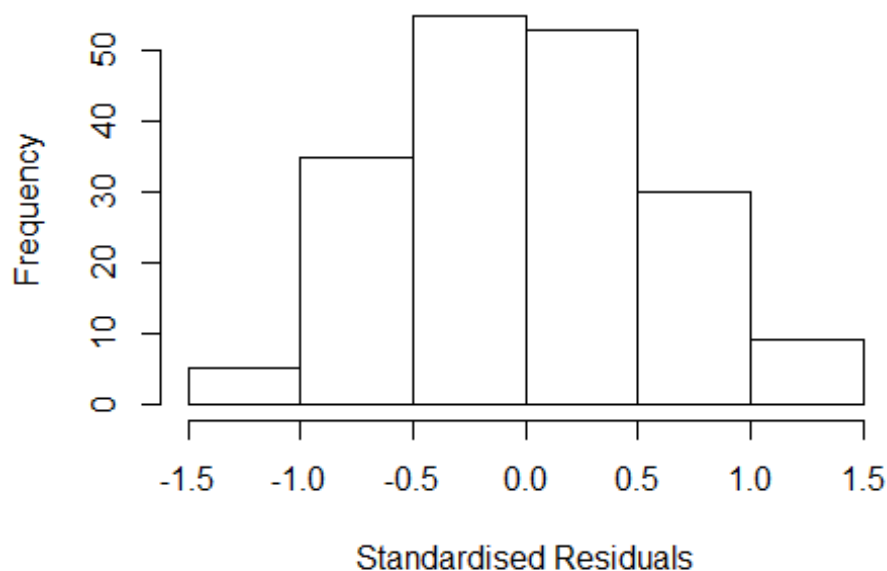
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.3654     0.2338   5.84 2.6e-08 ***
## LastAuthorFemale1  0.0608     0.1433   0.42  0.672
## Year1998         -0.1248     0.2933  -0.43  0.671
## Year1999         -0.5101     0.3003  -1.70  0.091 .
## Year2000         -0.5285     0.2807  -1.88  0.061 .
## Year2001         -0.5700     0.3509  -1.62  0.106
## Year2002         -0.3273     0.2824  -1.16  0.248
## Year2003         -0.1967     0.2977  -0.66  0.510
## Year2004         -0.2097     0.2692  -0.78  0.437
## Year2005         -0.0852     0.2908  -0.29  0.770
## Year2006         -0.5216     0.2638  -1.98  0.050 *
## Year2007          0.0761     0.3013   0.25  0.801
## Year2008         -0.5968     0.2886  -2.07  0.040 *
## Year2009         -0.2284     0.3134  -0.73  0.467
## Year2010         -0.4686     0.2480  -1.89  0.061 .
## Year2011         -0.2976     0.2993  -0.99  0.322
## Year2012         -0.0708     0.3156  -0.22  0.823
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.643
## Multiple R-squared:  0.1, Adjusted R-squared:  0.0156
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 168 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.595  0.896  0.950  0.924  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      5.35e-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: 187"
## [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
##    8    5    7   22    4    9    4    9   13    7   10   11    7    7    4
## 2011 2012
##   12    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    1    2    1    1    1    2    4    5    1    1    1    3    3    4
## 2011 2012
##    6    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    1    2    0    1    0    2    3    3    1    1    1    3    3    4
## 2011 2012
##    6    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 1.8612097182042"

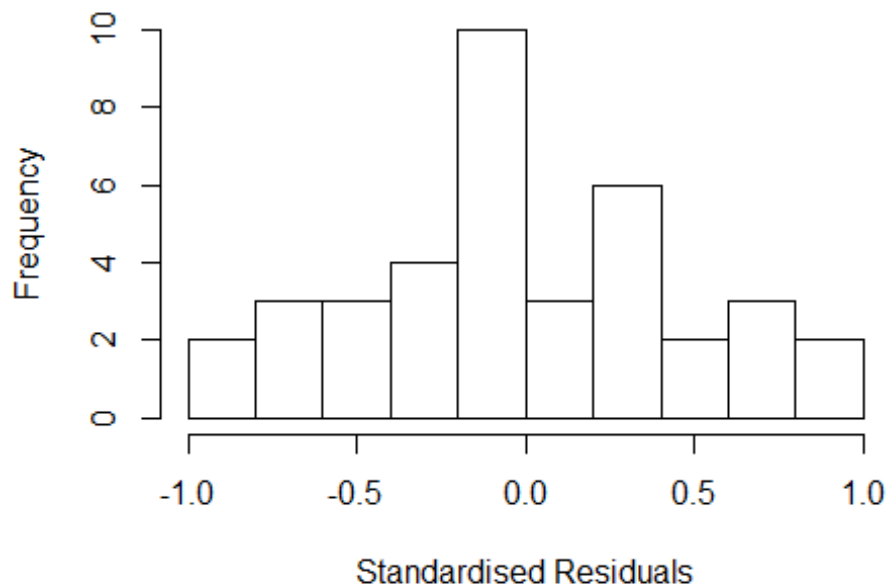
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 1.90636858599387"
## [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"
##
##          GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 0.7555 1      0.8692
## LastAuthorFemale 0.6274 1      0.7921
## Year              4.2037 14      1.0526
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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"
```

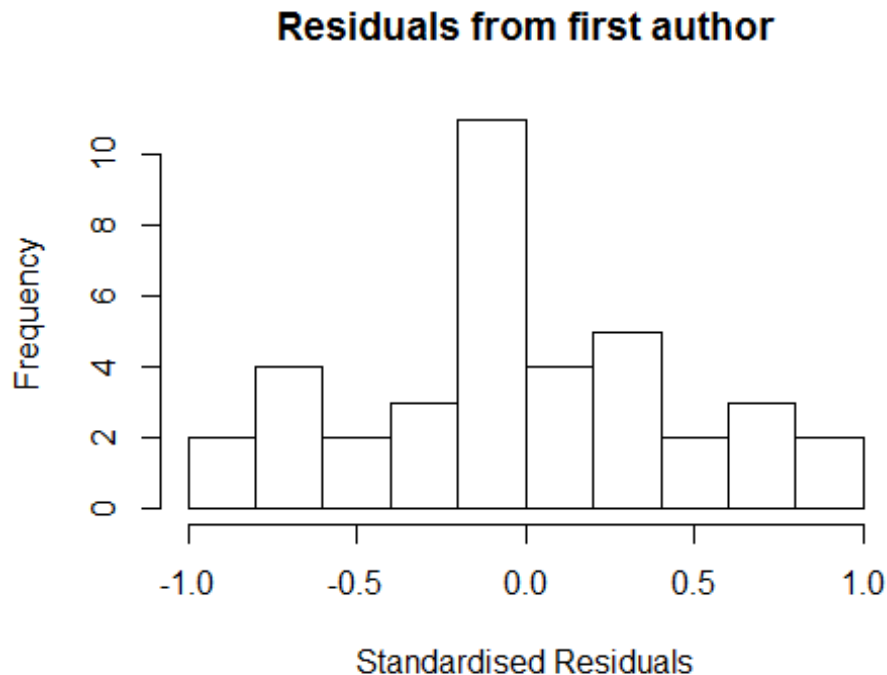


```

## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.94e-01 -3.08e-01 -3.05e-16  2.44e-01  8.90e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2279      0.3856   3.18  0.0045 **
## FirstAuthorFemale1 -0.5554      0.2688  -2.07  0.0514 .
## LastAuthorFemale1 -0.2689      0.2383  -1.13  0.2719
## Year1997          -0.5249      0.3856  -1.36  0.1879
## Year1998           0.4250      0.4308   0.99  0.3351
## Year2000          -0.4679      0.3856  -1.21  0.2384
## Year2002          -0.8589      0.4724  -1.82  0.0833 .
## Year2003           0.3980      0.6095   0.65  0.5208
## Year2004           0.5733      0.4224   1.36  0.1890
## Year2005          -0.3769      0.3856  -0.98  0.3395
## Year2006          -0.4429      0.3856  -1.15  0.2636
## Year2007          -0.4789      0.3856  -1.24  0.2279
## Year2008           0.3227      0.5206   0.62  0.5420
## Year2009           0.4359      0.6526   0.67  0.5114
## Year2010           0.1037      0.4635   0.22  0.8252
## Year2011           0.0274      0.4414   0.06  0.9511
## Year2012           0.0575      0.4059   0.14  0.8887
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.756
## Multiple R-squared:  0.377, Adjusted R-squared:  -0.0974
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  9 weights are ~= 1. The remaining 29 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.849  0.929  0.974   0.957   0.992   0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      2.63e-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 2.042  1          1.429
## Year              2.042 14          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 + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.96e-01 -2.54e-01  1.39e-16  2.45e-01  8.91e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.22703    0.38995   3.15  0.0047 **
## FirstAuthorFemale1 -0.48120    0.32496  -1.48  0.1528
## Year1997        -0.52403    0.38995  -1.34  0.1927
## Year1998         0.29147    0.39390   0.74  0.4671
## Year2000        -0.46703    0.38995  -1.20  0.2438
## Year2002        -0.85803    0.47673  -1.80  0.0856 .
## Year2003         0.39843    0.61828   0.64  0.5260
```

```

## Year2004          0.57455      0.42653      1.35      0.1917
## Year2005          -0.37603      0.38995     -0.96      0.3454
## Year2006          -0.44203      0.38995     -1.13      0.2692
## Year2007          -0.47803      0.38995     -1.23      0.2332
## Year2008           0.18531      0.58417      0.32      0.7541
## Year2009           0.43885      0.66389      0.66      0.5155
## Year2010           0.08505      0.47096      0.18      0.8583
## Year2011           0.01557      0.44821      0.03      0.9726
## Year2012          -0.00926      0.40050     -0.02      0.9818
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.719
## Multiple R-squared:  0.37,   Adjusted R-squared:  -0.0589
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  11 weights are ~= 1. The remaining 27 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.833  0.915  0.969   0.949   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      2.63e-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 3.669  1      1.915
## Year             3.669 14      1.048

## [1] "List of  0 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
## -9.93e-01 -2.84e-01 -6.94e-16  2.80e-01  8.90e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.2285      0.3827   3.21  0.004 **
## LastAuthorFemale1 -0.0832      0.1543  -0.54  0.595
## Year1997          -0.5255      0.3827  -1.37  0.184
## Year1998           0.3316      0.3994   0.83  0.415
## Year2000          -0.4685      0.3827  -1.22  0.234
## Year2002          -0.8595      0.4696  -1.83  0.081 .
## Year2003           0.3977      0.6041   0.66  0.517
## Year2004           0.5726      0.4195   1.37  0.186
## Year2005          -0.3775      0.3827  -0.99  0.335
## Year2006          -0.4435      0.3827  -1.16  0.259
## Year2007          -0.4795      0.3827  -1.25  0.223
## Year2008          -0.1073      0.4523  -0.24  0.815
## Year2009           0.4341      0.6449   0.67  0.508
## Year2010          -0.0437      0.4662  -0.09  0.926
## Year2011          -0.0588      0.4639  -0.13  0.900
## Year2012           0.0103      0.3984   0.03  0.980
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.782
## Multiple R-squared:  0.332, Adjusted R-squared:  -0.123
## Convergence in 10 IRWLS iterations
##
## Robustness weights:
##  9 weights are ~= 1. The remaining 29 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.859  0.929  0.975   0.957   0.991   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      2.63e-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: 38"

```

```

## [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
## 12 8 4 11 9 9 3 4 5 7 5 8 11 8 6
## 2011 2012
## 9 6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 1 1 5 3 2 0 0 3 6 4 5 8 5 3
## 2011 2012
## 4 5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 1 0 1 5 2 2 0 0 3 6 4 5 7 4 3
## 2011 2012
## 4 5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 1.5874010519682"

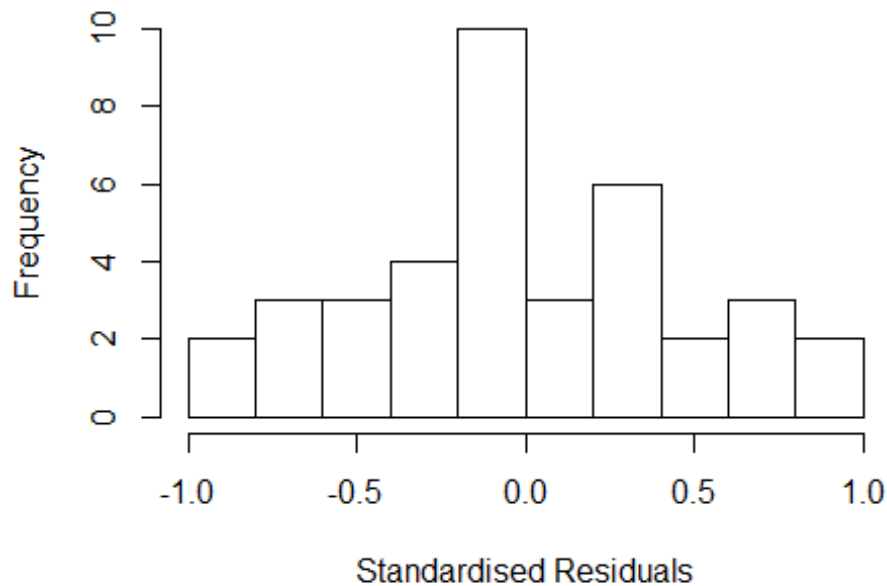
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.3
## 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: 1.81712059283214"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1.5, 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"
## [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 80.2  1           8.956
## Year              80.2 13           1.184
##
## [1] "List of 0 outliers 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
```

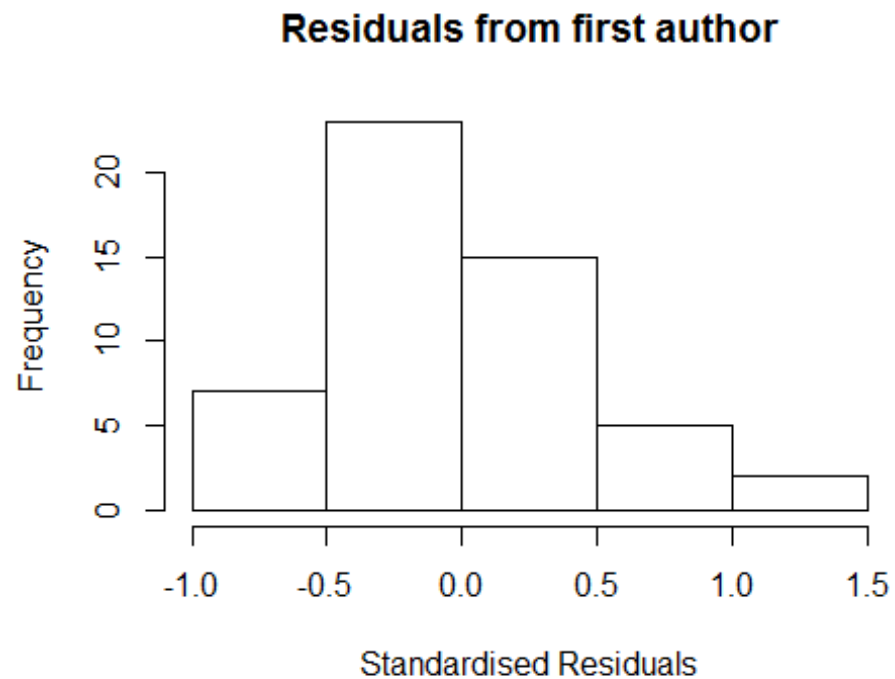
```

## -0.9680 -0.2725 -0.0219  0.3142  1.4200
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    8.36e-01   1.79e-08  4.68e+07  <2e-16 ***
## FirstAuthorFemale1 4.15e-02   2.22e-01   1.90e-01   0.8528
## Year1998        -1.55e-01   2.43e-08 -6.38e+06  <2e-16 ***
## Year1999        -1.17e-01   6.28e-02 -1.86e+00   0.0705 .
## Year2000         3.14e-01   1.12e-01   2.81e+00   0.0078 **
## Year2001         4.36e-01   3.35e-01   1.30e+00   0.2007
## Year2004         1.64e-01   1.98e-01   8.30e-01   0.4141
## Year2005         2.82e-02   2.07e-01   1.40e-01   0.8926
## Year2006        -4.19e-01   2.36e-01 -1.77e+00   0.0847 .
## Year2007        -3.36e-01   2.65e-01 -1.27e+00   0.2128
## Year2008         1.32e-01   2.35e-01   5.60e-01   0.5776
## Year2009        -1.29e-01   5.12e-01 -2.50e-01   0.8032
## Year2010         3.99e-02   1.38e-01   2.90e-01   0.7732
## Year2011        -1.89e-02   1.39e-01 -1.40e-01   0.8926
## Year2012         2.94e-01   3.37e-01   8.70e-01   0.3881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.198, Adjusted R-squared:  -0.105
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~ = 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.379  0.887  0.944  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      1.92e-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

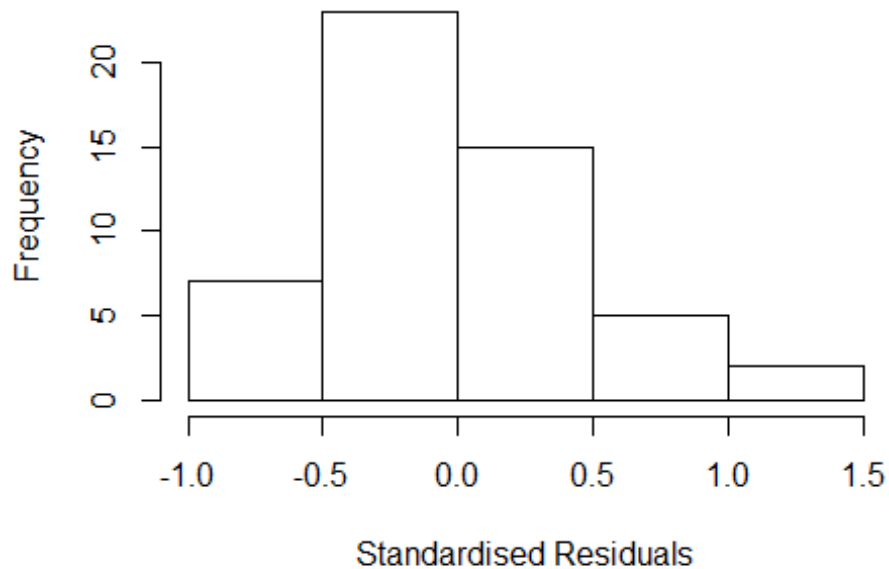
```

```
## 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 13          NaN
```


Residuals from last author



```
## [1] "List of 0 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
## -0.9680 -0.2725 -0.0219 0.3142 1.4200
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8360 0.0000 Inf <2e-16 ***
## LastAuthorFemale1 0.0415 0.2220 0.19 0.8528
## Year1998 -0.1550 0.0000 -Inf <2e-16 ***
## Year1999 -0.1170 0.0628 -1.86 0.0705 .
## Year2000 0.3143 0.1117 2.81 0.0078 **
## Year2001 0.4358 0.3256 1.34 0.1889
## Year2004 0.1637 0.1982 0.83 0.4141
## Year2005 0.0282 0.2072 0.14 0.8926
## Year2006 -0.4187 0.2363 -1.77 0.0847 .
## Year2007 -0.3360 0.2651 -1.27 0.2128
## Year2008 0.1320 0.2350 0.56 0.5776
## Year2009 -0.1286 0.5125 -0.25 0.8032
```

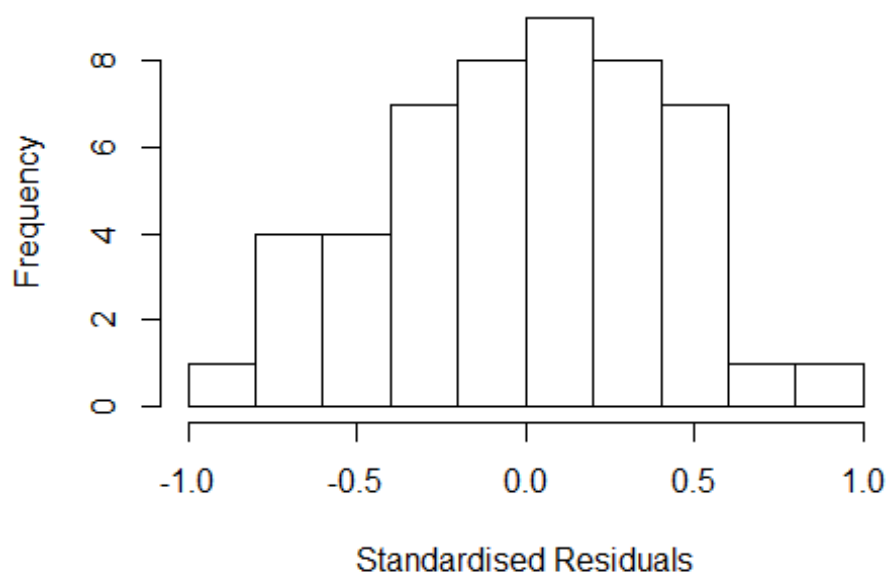
```

## Year2010          0.0399      0.1375      0.29      0.7732
## Year2011         -0.0189      0.1387     -0.14      0.8926
## Year2012          0.2945      0.3372      0.87      0.3881
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.198, Adjusted R-squared:  -0.105
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.379  0.887  0.944  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      1.92e-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: 52"
## [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
##    6    9    7    8   18    8    7    5    7   13    6    4    8    7    9
## 2011 2012
##   10   11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    6    3    3    7    4    2    2    2    4    2    2    4    2    1
## 2011 2012
##    3    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```
##      2      6      3      3      7      4      2      1      2      3      2      2      4      2      1
## 2011 2012
##      3      3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: NaN"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.376e+15  1      3.709e+07
## LastAuthorFemale  7.484e+14  1      2.736e+07
## UniqueAuthors    1.098e+30  3      1.016e+05
## Year              4.152e+33 16      1.123e+01
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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:
```

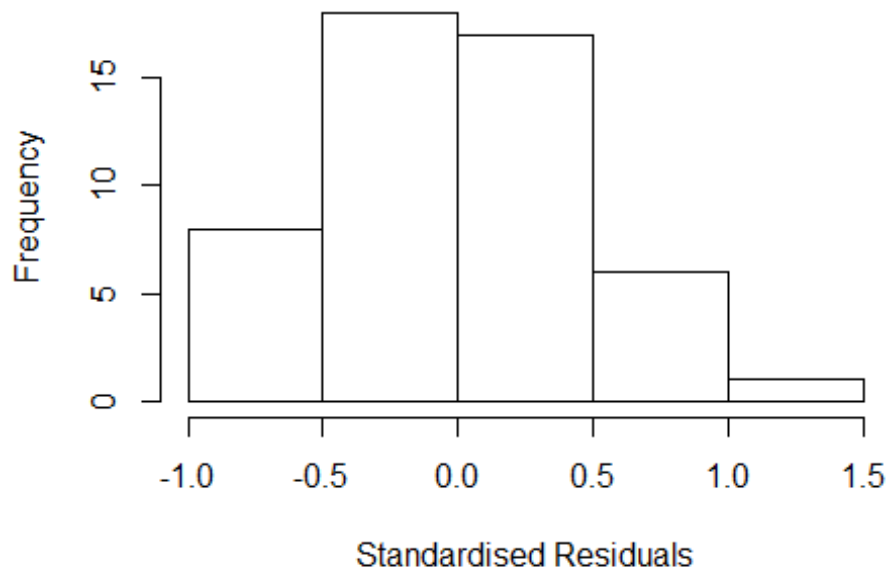
```

##      Min      1Q  Median      3Q      Max
## -0.8204 -0.2631  0.0114  0.2539  0.8138
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9427     0.3322   2.84  0.0084 **
## FirstAuthorFemale1 -0.0868     0.4423  -0.20  0.8458
## LastAuthorFemale1  0.0615     0.3178   0.19  0.8480
## UniqueAuthors2     0.4616     0.1779   2.60  0.0149 *
## UniqueAuthors3     0.2270     0.2864   0.79  0.4347
## UniqueAuthors5    -1.0285     0.3789  -2.71  0.0112 *
## Year1997           0.1112     0.4017   0.28  0.7840
## Year1998           0.3424     0.4607   0.74  0.4635
## Year1999          -0.2186     0.3501  -0.62  0.5374
## Year2000          -0.2593     0.3896  -0.67  0.5110
## Year2001          -0.1387     0.3612  -0.38  0.7039
## Year2002           0.2745     0.5843   0.47  0.6421
## Year2003           0.9313     0.3322   2.80  0.0091 **
## Year2004          -0.0172     0.3888  -0.04  0.9650
## Year2005           0.1656     0.3774   0.44  0.6643
## Year2006          -0.5883     0.3875  -1.52  0.1401
## Year2007          -0.6282     0.4074  -1.54  0.1344
## Year2008          -0.4971     0.4817  -1.03  0.3109
## Year2009          -0.7227     0.3687  -1.96  0.0600 .
## Year2010          -0.1213     0.3346  -0.36  0.7197
## Year2011          -0.0626     0.3371  -0.19  0.8541
## Year2012          -0.3194     0.5812  -0.55  0.5870
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.564
## Multiple R-squared:  0.486, Adjusted R-squared:  0.101
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 44 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.817  0.932  0.975  0.953  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          2.00e-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 9864.96 1          99.323
## LastAuthorFemale  17.81 1          4.220
## Year              19970.01 16          1.363
```

Residuals from first and last author



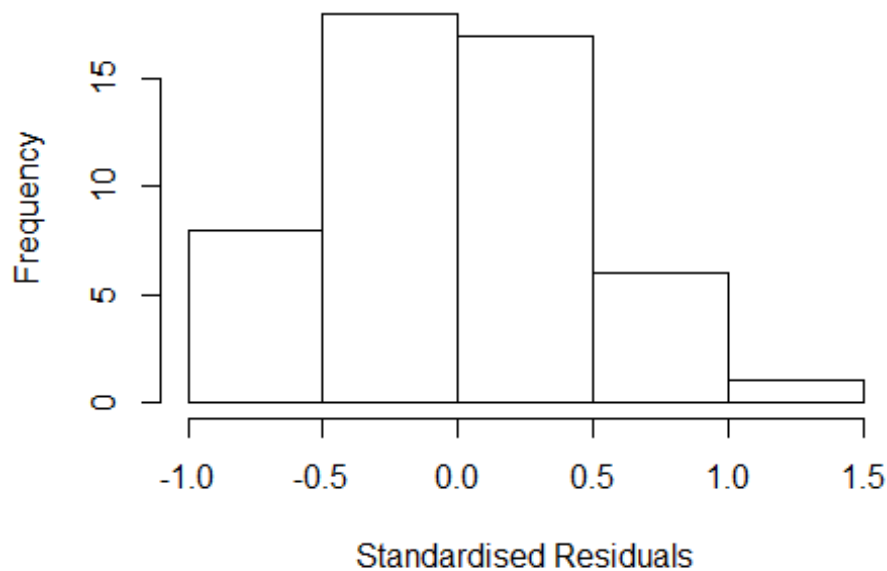
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.35e-01 -2.97e-01 2.64e-16 3.12e-01 1.18e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1735 0.5879 2.00 0.055 .
```

```

## FirstAuthorFemale1 -0.3118      0.4913     -0.63      0.530
## LastAuthorFemale1  0.0567      0.2948      0.19      0.849
## Year1997           -0.1053      0.6456     -0.16      0.872
## Year1998            0.2432      0.6761      0.36      0.721
## Year1999           -0.3011      0.6402     -0.47      0.641
## Year2000           -0.3543      0.6461     -0.55      0.587
## Year2001           -0.0867      0.6119     -0.14      0.888
## Year2002            0.7303      0.7765      0.94      0.354
## Year2003            0.7005      0.5879      1.19      0.243
## Year2004           -0.1345      0.6394     -0.21      0.835
## Year2005           -0.0649      0.6138     -0.11      0.917
## Year2006           -0.3575      0.6193     -0.58      0.568
## Year2007           -0.8590      0.6331     -1.36      0.185
## Year2008           -0.4496      0.6181     -0.73      0.472
## Year2009           -0.9535      0.6092     -1.57      0.128
## Year2010            0.1095      0.5879      0.19      0.853
## Year2011            0.0172      0.6074      0.03      0.978
## Year2012           -0.3384      0.7593     -0.45      0.659
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.342, Adjusted R-squared:  -0.0404
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.667  0.916  0.971  0.942  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      2.00e-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 1938 1      44.022
## Year               1938 16      1.267

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -8.63e-01 -3.08e-01  1.87e-16  3.10e-01  1.14e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1735     0.5999   1.96   0.059 .
## FirstAuthorFemale1 -0.2774     0.4051  -0.68   0.498
## Year1997         -0.1062     0.6544  -0.16   0.872
## Year1998          0.2487     0.6925   0.36   0.722
## Year1999         -0.3013     0.6517  -0.46   0.647
## Year2000         -0.3498     0.6586  -0.53   0.599
## Year2001         -0.0811     0.6248  -0.13   0.898
## Year2002          0.6959     0.7302   0.95   0.348
## Year2003          0.7005     0.5999   1.17   0.252
## Year2004         -0.1345     0.6508  -0.21   0.838
## Year2005         -0.0652     0.6254  -0.10   0.918
## Year2006         -0.3575     0.6309  -0.57   0.575
```

```

## Year2007          -0.8590      0.6445   -1.33    0.192
## Year2008          -0.4423      0.6294   -0.70    0.487
## Year2009          -0.9535      0.6209   -1.54    0.134
## Year2010           0.1095      0.5999    0.18    0.856
## Year2011           0.0361      0.6171    0.06    0.954
## Year2012          -0.3100      0.7360   -0.42    0.676
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.569
## Multiple R-squared:  0.342, Adjusted R-squared:  -0.00809
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 43 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.666  0.914  0.969  0.938  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      2.00e-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 2.166 1      1.472
## Year              2.166 16      1.024

## [1] "List of 0 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
## -9.35e-01 -3.02e-01 -3.05e-16  3.32e-01  8.99e-01

```



```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1735    0.5665   2.07  0.046 *
## LastAuthorFemale1 -0.0986    0.2756  -0.36  0.723
## Year1997          -0.1399    0.6282  -0.22  0.825
## Year1998           0.1980    0.6351   0.31  0.757
## Year1999          -0.3006    0.6202  -0.48  0.631
## Year2000          -0.3778    0.6231  -0.61  0.549
## Year2001          -0.1267    0.5845  -0.22  0.830
## Year2002           0.4185    0.5666   0.74  0.465
## Year2003           0.7005    0.5665   1.24  0.225
## Year2004          -0.1345    0.6189  -0.22  0.829
## Year2005          -0.0643    0.5940  -0.11  0.915
## Year2006          -0.3575    0.5987  -0.60  0.555
## Year2007          -0.8590    0.6126  -1.40  0.170
## Year2008          -0.4747    0.6212  -0.76  0.450
## Year2009          -0.9535    0.5884  -1.62  0.115
## Year2010           0.1095    0.5665   0.19  0.848
## Year2011           0.0682    0.5966   0.11  0.910
## Year2012          -0.2762    0.7234  -0.38  0.705
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.626
## Multiple R-squared:  0.321, Adjusted R-squared:  -0.04
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 42 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.807  0.929   0.965   0.947   0.984   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           2.00e-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: 50"
## [1] ""

```

```

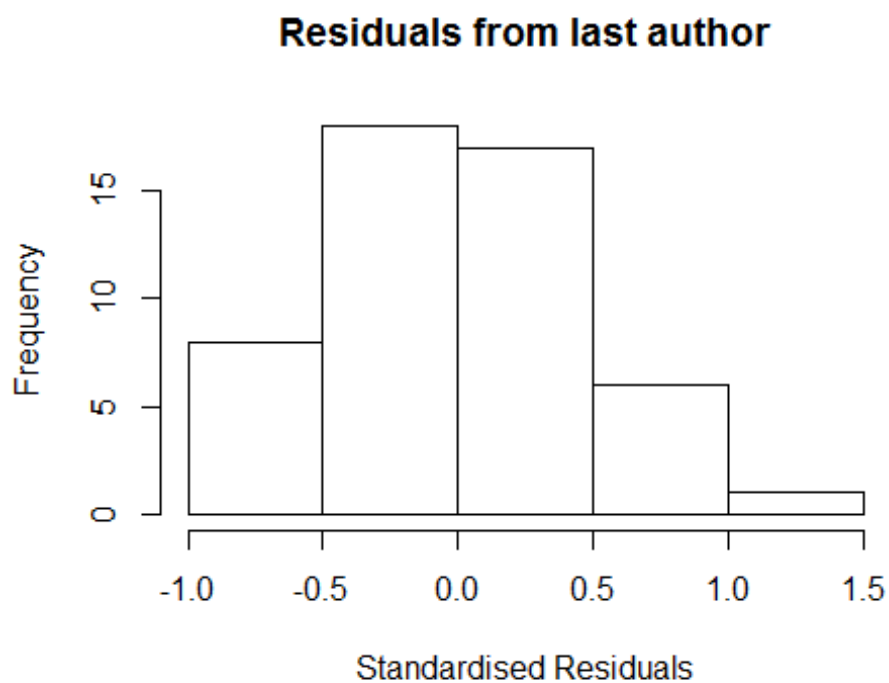
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3104"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    35    51    47     5    42    50    43    40    59    59    89    83   107    99    83
## 2011 2012
##    63    59
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     9    19    13     1     9    14    11    23    26    28    43    41    50    55    44
## 2011 2012
##    39    35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##     6    15    11     1     9    13     6    18    21    25    33    35    42    44    36
## 2011 2012
##    36    29
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.22948505376226"
## [1] "Male first author team size 2018 geometric mean: 3.2349560865714"

## 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 = 46, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5.47722557505166"
## [1] "Male last author team size 2018 geometric mean: 3.23764262397294"

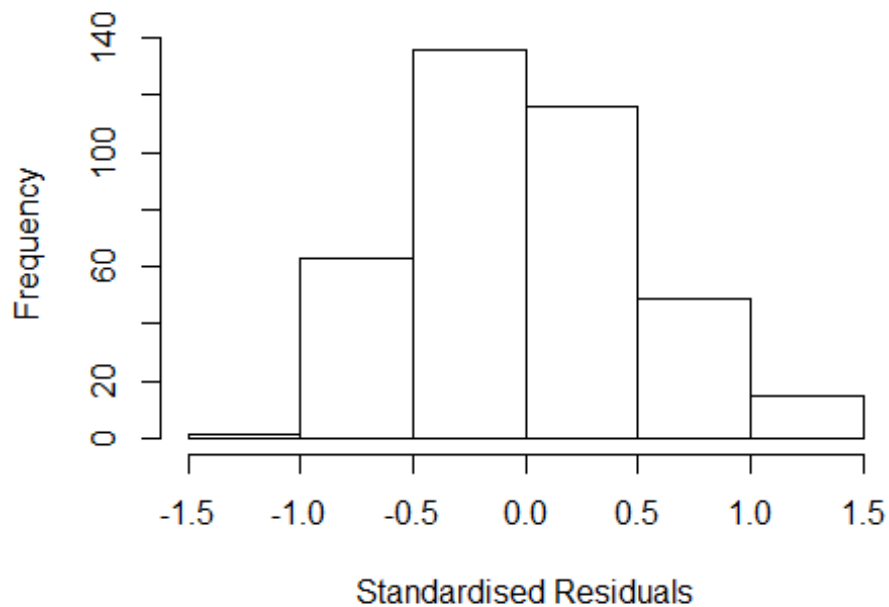
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 39, 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 1.789 1      1.338
## LastAuthorFemale  1.689 1      1.300
## UniqueAuthors    6.002 4      1.251
## Year              8.396 16     1.069
```

Residuals from first and last author and team size



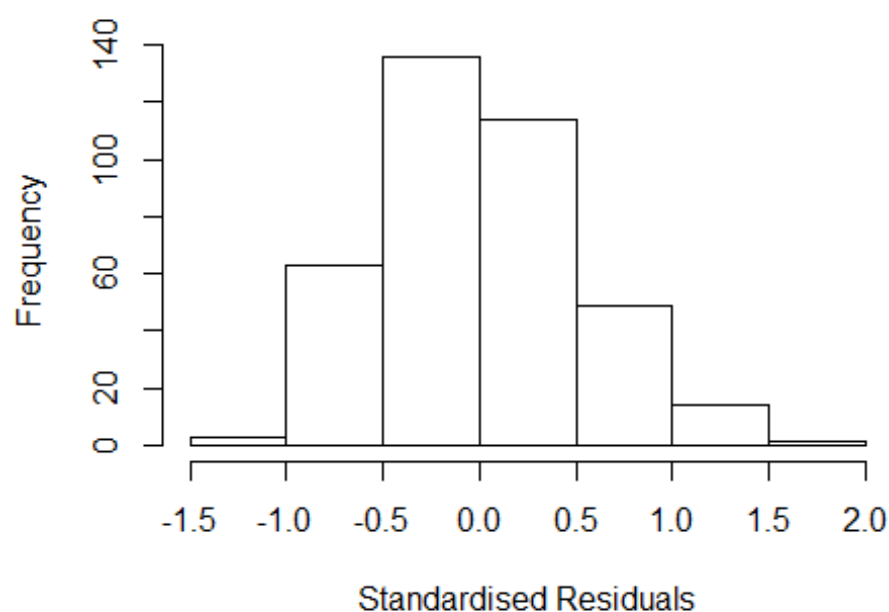
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.416 -0.345 -0.032 0.384 1.407
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5986 0.1712 9.34 < 2e-16 ***
## FirstAuthorFemale1 0.2091 0.0797 2.62 0.00905 **
## LastAuthorFemale1 -0.0349 0.0985 -0.35 0.72341
## UniqueAuthors2 -0.1892 0.0850 -2.23 0.02659 *
## UniqueAuthors3 -0.3023 0.0893 -3.39 0.00079 ***
## UniqueAuthors4 -0.2215 0.1096 -2.02 0.04397 *
## UniqueAuthors5 -0.1339 0.1323 -1.01 0.31188
## Year1997 -0.1993 0.2108 -0.95 0.34514
## Year1998 -0.3739 0.2443 -1.53 0.12681
## Year1999 0.4936 0.1730 2.85 0.00458 **
```

```

## Year2000          -0.3564      0.1950    -1.83    0.06844 .
## Year2001          -0.4542      0.2466    -1.84    0.06632 .
## Year2002           0.0389      0.3228     0.12    0.90413
## Year2003          -0.4333      0.2233    -1.94    0.05307 .
## Year2004          -0.3156      0.2064    -1.53    0.12721
## Year2005          -0.5026      0.1957    -2.57    0.01061 *
## Year2006          -0.8775      0.1835    -4.78    2.5e-06 ***
## Year2007          -0.7817      0.2006    -3.90    0.00012 ***
## Year2008          -0.6800      0.1860    -3.66    0.00030 ***
## Year2009          -0.6749      0.1970    -3.43    0.00069 ***
## Year2010          -0.4765      0.1995    -2.39    0.01746 *
## Year2011          -0.5445      0.1844    -2.95    0.00335 **
## Year2012          -0.4958      0.1977    -2.51    0.01260 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.541
## Multiple R-squared:  0.188, Adjusted R-squared:  0.138
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 351 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.473  0.884  0.951  0.917  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          2.63e-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 1.690 1          1.300
## LastAuthorFemale  1.415 1          1.189
## Year              1.829 16          1.019

```

Residuals from first and last author



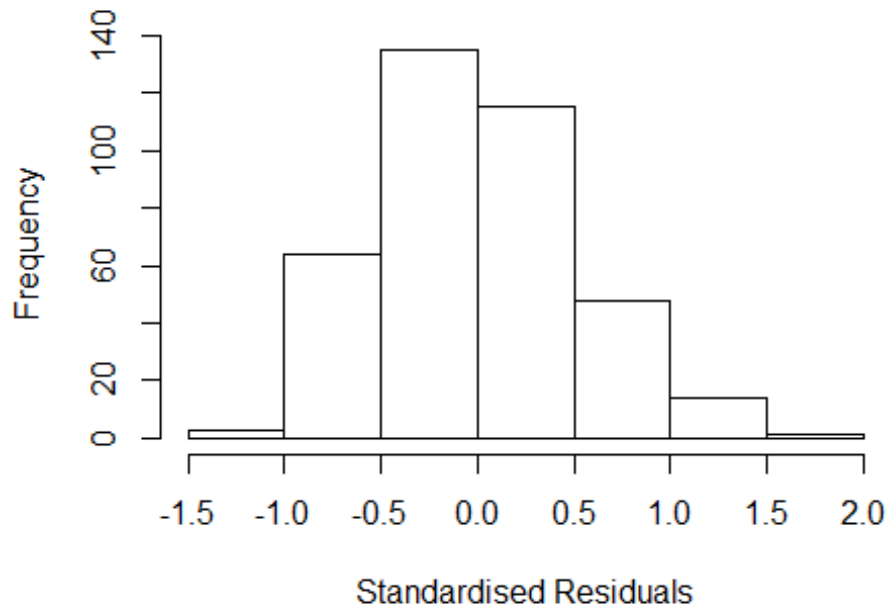
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4813 -0.4039 -0.0308 0.3610 1.5915
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4679 0.1669 8.80 < 2e-16 ***
## FirstAuthorFemale1 0.2008 0.0808 2.48 0.01343 *
## LastAuthorFemale1 -0.0345 0.1000 -0.35 0.72996
## Year1997 -0.2038 0.2144 -0.95 0.34259
## Year1998 -0.3512 0.2461 -1.43 0.15430
## Year1999 0.4351 0.1669 2.61 0.00950 **
## Year2000 -0.3922 0.1930 -2.03 0.04294 *
## Year2001 -0.5300 0.2364 -2.24 0.02557 *
## Year2002 0.0134 0.3041 0.04 0.96482
## Year2003 -0.4999 0.2219 -2.25 0.02489 *
## Year2004 -0.3856 0.2049 -1.88 0.06069 .
## Year2005 -0.5921 0.1929 -3.07 0.00231 **
```

```

## Year2006          -0.9571      0.1826   -5.24  2.7e-07 ***
## Year2007          -0.8288      0.2040   -4.06  5.9e-05 ***
## Year2008          -0.7304      0.1856   -3.94  1.0e-04 ***
## Year2009          -0.7424      0.1927   -3.85  0.00014 ***
## Year2010          -0.5195      0.1996   -2.60  0.00963 **
## Year2011          -0.5761      0.1852   -3.11  0.00201 **
## Year2012          -0.5288      0.1973   -2.68  0.00770 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.548
## Multiple R-squared:  0.16,   Adjusted R-squared:  0.118
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 342 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.380  0.880  0.950  0.914  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      2.63e-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 1.555 1          1.247
## Year              1.555 16          1.014

```

Residuals from first author



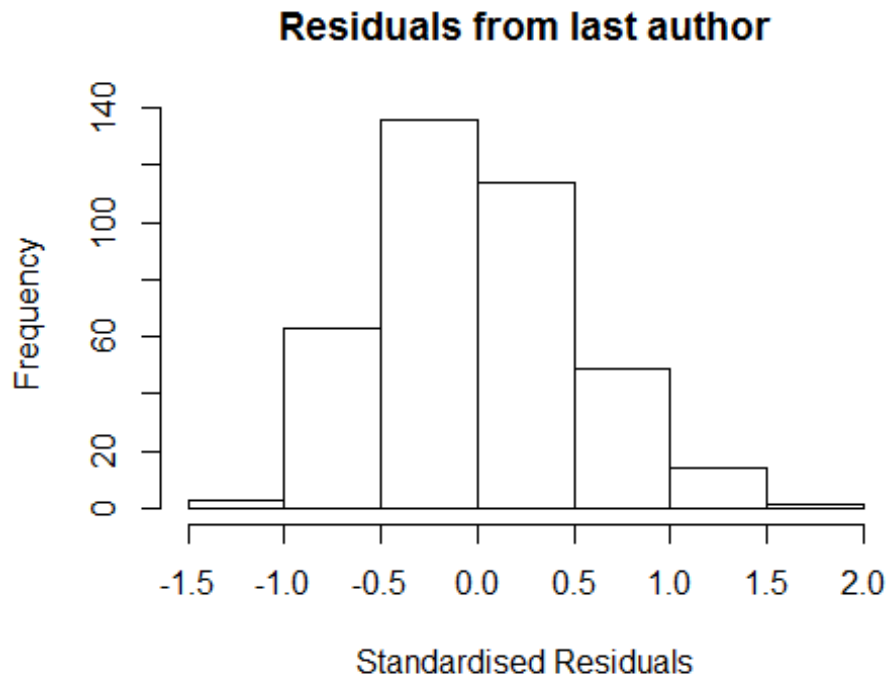
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4800 -0.4082 -0.0295 0.3603 1.5933
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4679 0.1668 8.80 < 2e-16 ***
## FirstAuthorFemale1 0.1937 0.0773 2.51 0.01262 *
## Year1997 -0.2058 0.2142 -0.96 0.33735
## Year1998 -0.3548 0.2455 -1.45 0.14923
## Year1999 0.4351 0.1668 2.61 0.00946 **
## Year2000 -0.3987 0.1923 -2.07 0.03884 *
## Year2001 -0.5376 0.2379 -2.26 0.02442 *
## Year2002 0.0121 0.3045 0.04 0.96844
## Year2003 -0.5043 0.2206 -2.29 0.02287 *
## Year2004 -0.3854 0.2046 -1.88 0.06039 .
## Year2005 -0.5996 0.1919 -3.12 0.00193 **
## Year2006 -0.9592 0.1826 -5.25 2.6e-07 ***
```



```

## Year2007          -0.8295      0.2033   -4.08  5.5e-05 ***
## Year2008          -0.7340      0.1854   -3.96  9.1e-05 ***
## Year2009          -0.7442      0.1924   -3.87  0.00013 ***
## Year2010          -0.5249      0.1992   -2.63  0.00878 **
## Year2011          -0.5772      0.1850   -3.12  0.00195 **
## Year2012          -0.5285      0.1971   -2.68  0.00769 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.551
## Multiple R-squared:  0.159, Adjusted R-squared:  0.12
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 344 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.384  0.882  0.948  0.916  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      2.63e-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 1.236 1          1.112
## Year              1.236 16          1.007

```



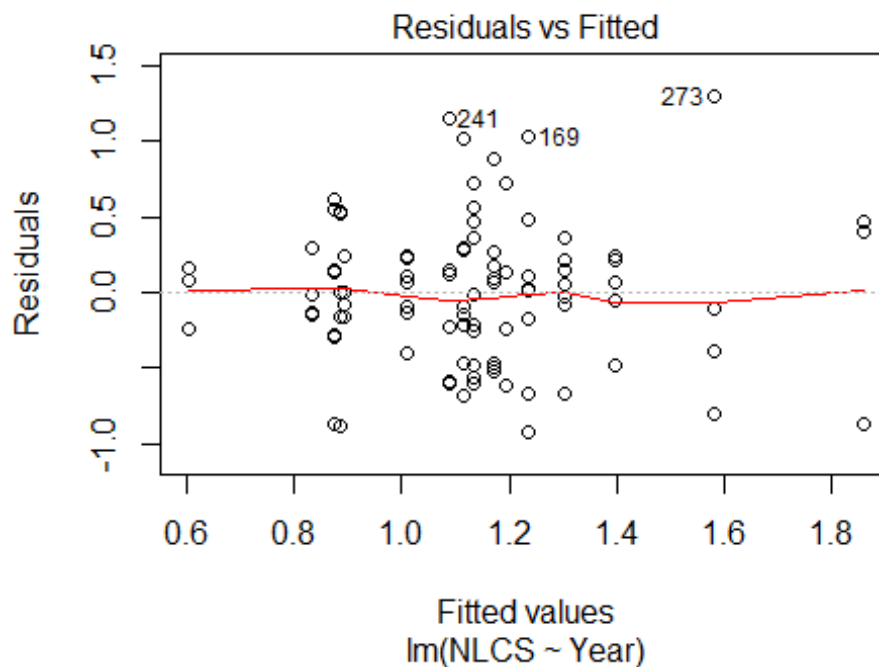
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4768 -0.4253 -0.0383 0.3589 1.5601
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.46807 0.16647 8.82 < 2e-16 ***
## LastAuthorFemale1 0.02080 0.09909 0.21 0.83385
## Year1997 -0.15300 0.21944 -0.70 0.48611
## Year1998 -0.33498 0.24458 -1.37 0.17166
## Year1999 0.43493 0.16647 2.61 0.00936 **
## Year2000 -0.38550 0.19822 -1.94 0.05257 .
## Year2001 -0.52005 0.23556 -2.21 0.02789 *
## Year2002 0.00875 0.30586 0.03 0.97719
## Year2003 -0.51226 0.22294 -2.30 0.02215 *
## Year2004 -0.35611 0.20883 -1.71 0.08900 .
## Year2005 -0.53744 0.19580 -2.74 0.00636 **
## Year2006 -0.93278 0.18470 -5.05 7.0e-07 ***
```

```

## Year2007          -0.80514      0.20025    -4.02   7.1e-05 ***
## Year2008          -0.70862      0.18360    -3.86   0.00013 ***
## Year2009          -0.71113      0.19104    -3.72   0.00023 ***
## Year2010          -0.49105      0.19733    -2.49   0.01328 *
## Year2011          -0.54978      0.18399    -2.99   0.00300 **
## Year2012          -0.47185      0.19620    -2.40   0.01668 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.559
## Multiple R-squared:  0.142, Adjusted R-squared:  0.102
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 353 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.416  0.885   0.948   0.919   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      2.63e-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: 380"
## [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
##   10   13   6   20   18   12   11   11   12   7   10   18   12   12   16
## 2011 2012
##   20   11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    4    3    3    8    5    4    7    5    4    5    9    7    7    6
## 2011 2012

```

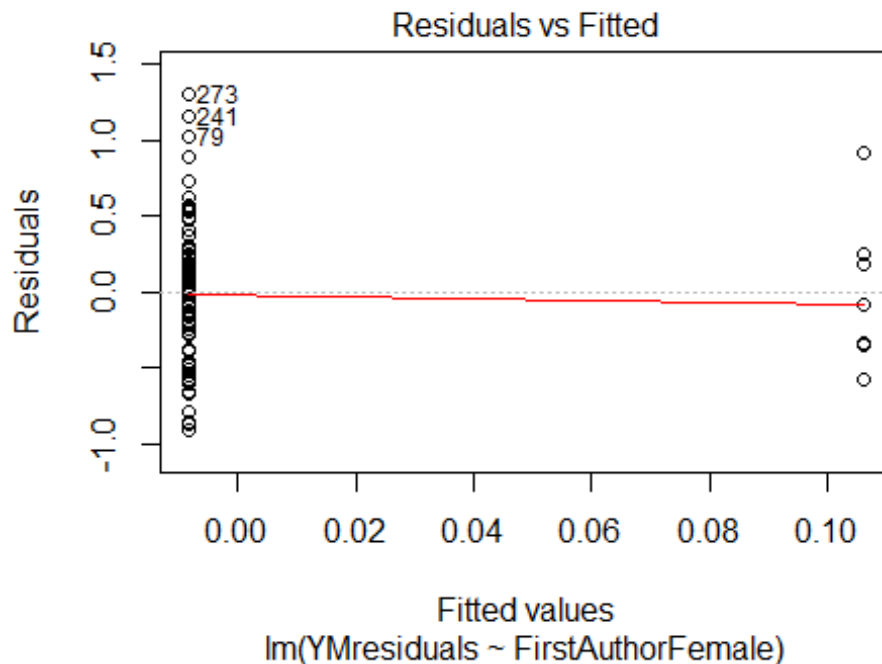
```
##      8      4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      4      3      2      2      7      4      4      5      4      4      4      8      7      5      6
## 2011 2012
##      5      3
## [1] "Heteroscedasticity checks, confirming that there are 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 = 0.094, df = 1, p-value = 0.8
## [1] "Female first author team size 2018 geometric mean: 2.71080601082953"
## [1] "Male first author team size 2018 geometric mean: 3.51991084883033"
## 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 = 15, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.88449914061482"
## [1] "Male last author team size 2018 geometric mean: 3.3658654363386"

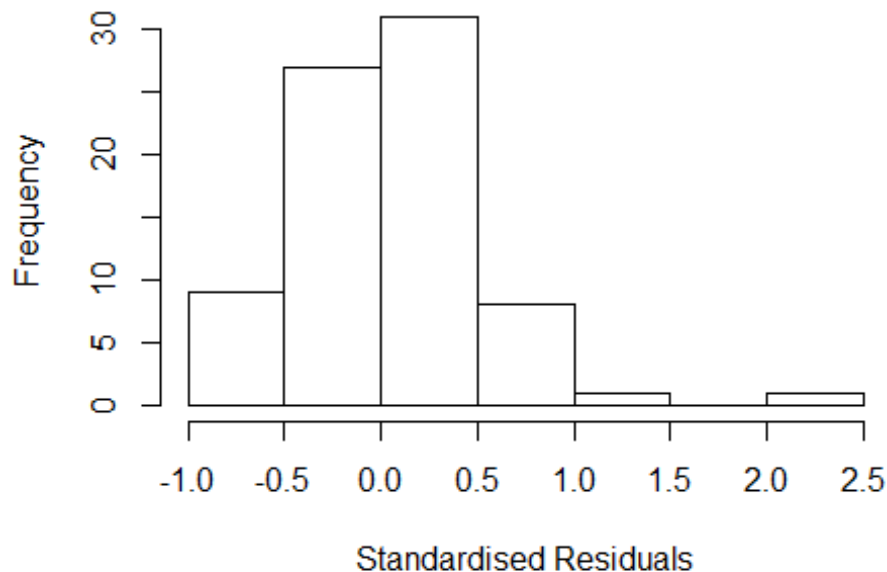
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16, 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	5.220	1	2.285
LastAuthorFemale	7.074	1	2.660
UniqueAuthors	122.484	4	1.824
Year	449.671	16	1.210

Residuals from first and last author and team size



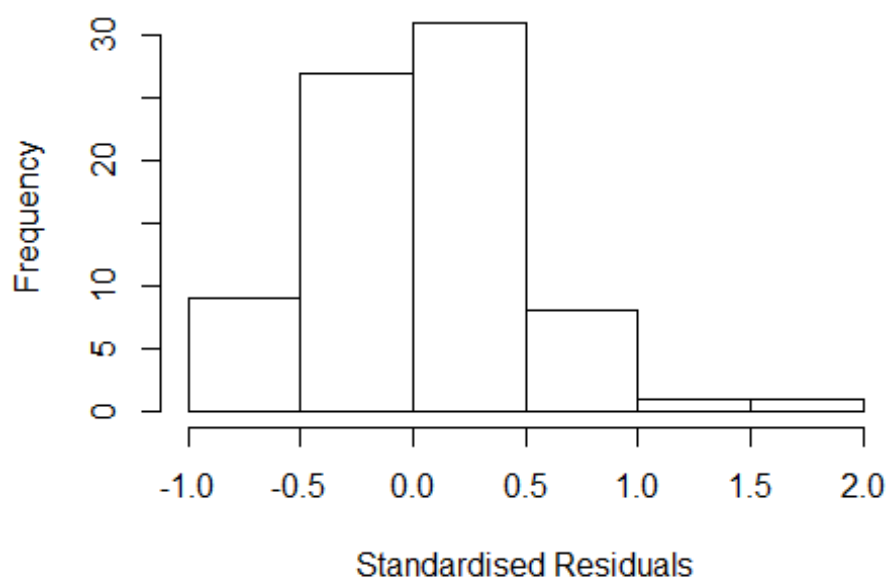
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8467 -0.2129 0.0109 0.1998 2.0291
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1469 0.3235 3.55 0.00082 ***
## FirstAuthorFemale1 0.4913 0.2855 1.72 0.09099 .
## LastAuthorFemale1 -0.5847 0.2475 -2.36 0.02179 *
## UniqueAuthors2 -0.1419 0.2204 -0.64 0.52245
## UniqueAuthors3 -0.1794 0.1811 -0.99 0.32633
## UniqueAuthors4 -0.3993 0.1703 -2.35 0.02272 *
## UniqueAuthors5 -0.2143 0.2676 -0.80 0.42664
## Year1997 -0.0767 0.2723 -0.28 0.77926
## Year1998 -0.1568 0.2743 -0.57 0.57001
## Year1999 0.6168 0.6619 0.93 0.35560
```

```

## Year2000          -0.0162      0.3257   -0.05  0.96063
## Year2001           0.1066      0.3134    0.34  0.73500
## Year2002          -0.1538      0.3149   -0.49  0.62719
## Year2003           0.3021      0.3123    0.97  0.33757
## Year2004           0.3132      0.3405    0.92  0.36171
## Year2005           0.1816      0.4483    0.41  0.68706
## Year2006           0.1708      0.3507    0.49  0.62826
## Year2007           0.2157      0.3269    0.66  0.51218
## Year2008           0.1052      0.2739    0.38  0.70229
## Year2009          -0.1583      0.4385   -0.36  0.71955
## Year2010           0.0928      0.3311    0.28  0.78029
## Year2011           0.1572      0.4246    0.37  0.71271
## Year2012          -0.1601      0.3657   -0.44  0.66323
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.437
## Multiple R-squared:  0.282, Adjusted R-squared:  -0.011
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## observation 75 is an outlier with |weight| <= 0.00024 ( < 0.0013);
## 8 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.452  0.872  0.978  0.924  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           1.30e-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 7.012 1 2.648
## LastAuthorFemale 8.824 1 2.971
## Year 6.718 16 1.061

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -0.8398 -0.2270 0.0283 0.2356 1.8199
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0347 0.2028 5.10 3.9e-06 ***
## FirstAuthorFemale1 0.3960 0.3641 1.09 0.281
## LastAuthorFemale1 -0.5532 0.2672 -2.07 0.043 *
## Year1997 -0.1157 0.2480 -0.47 0.643
## Year1998 -0.3152 0.2047 -1.54 0.129
## Year1999 0.6218 0.8635 0.72 0.474
## Year2000 -0.0390 0.2435 -0.16 0.873
## Year2001 0.0756 0.2649 0.29 0.776
## Year2002 -0.1427 0.2160 -0.66 0.512
## Year2003 0.2345 0.2488 0.94 0.350
## Year2004 0.3881 0.2539 1.53 0.132
## Year2005 0.1929 0.3818 0.51 0.615
```

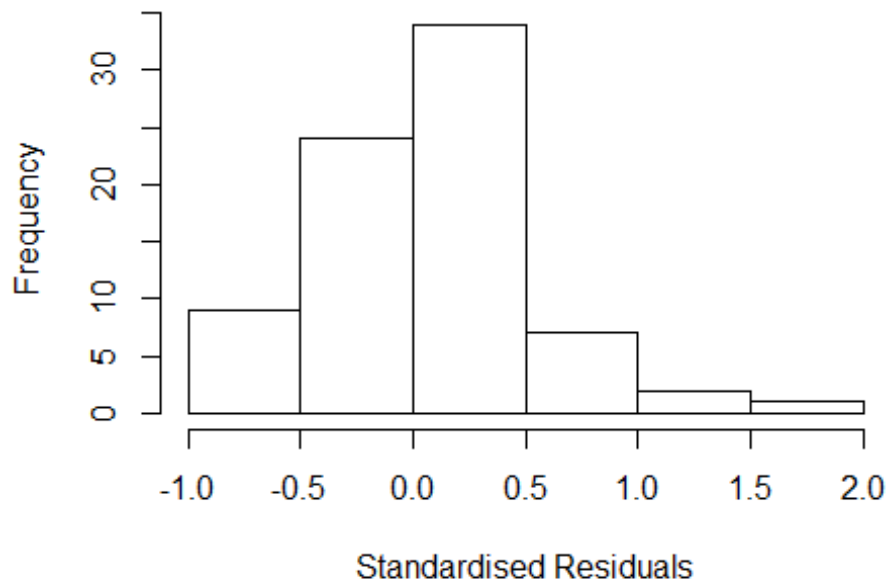


```

## Year2006          0.0611      0.3592      0.17      0.866
## Year2007          0.0949      0.2932      0.32      0.747
## Year2008         -0.0213      0.2187     -0.10      0.923
## Year2009         -0.1948      0.3154     -0.62      0.539
## Year2010         -0.0516      0.3072     -0.17      0.867
## Year2011          0.1552      0.3394      0.46      0.649
## Year2012          0.0194      0.3659      0.05      0.958
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.453
## Multiple R-squared:  0.215, Adjusted R-squared:  -0.0291
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 70 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0698 0.8760 0.9720 0.9130 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      1.30e-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 1.587 1      1.260
## Year              1.587 16      1.015

```

Residuals from first author



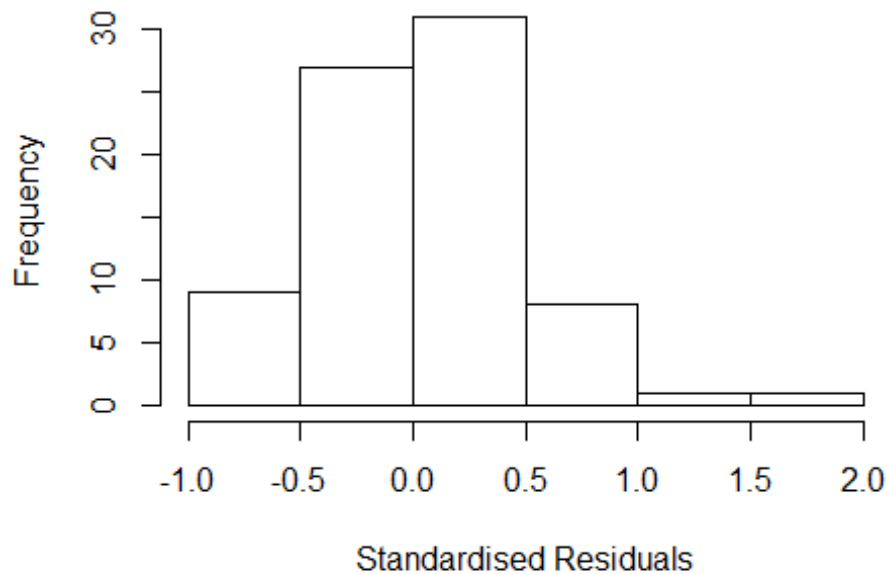
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8801 -0.2641 0.0385 0.2196 1.8212
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0346 0.2029 5.10 3.8e-06 ***
## FirstAuthorFemale1 0.0569 0.3478 0.16 0.87
## Year1997 -0.2967 0.2060 -1.44 0.16
## Year1998 -0.3151 0.2047 -1.54 0.13
## Year1999 0.6219 0.8636 0.72 0.47
## Year2000 -0.0705 0.2493 -0.28 0.78
## Year2001 0.0756 0.2650 0.29 0.78
## Year2002 -0.1426 0.2160 -0.66 0.51
## Year2003 0.2346 0.2488 0.94 0.35
## Year2004 0.3881 0.2539 1.53 0.13
## Year2005 0.1209 0.4009 0.30 0.76
## Year2006 -0.1165 0.3433 -0.34 0.74
```

```

## Year2007          0.1645      0.2972      0.55      0.58
## Year2008         -0.0212      0.2187     -0.10      0.92
## Year2009         -0.1947      0.3154     -0.62      0.54
## Year2010         -0.0518      0.3072     -0.17      0.87
## Year2011          0.1551      0.3395      0.46      0.65
## Year2012          0.0181      0.3650      0.05      0.96
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.452
## Multiple R-squared:  0.175, Adjusted R-squared:  -0.063
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0684 0.8850 0.9710 0.9050 0.9900 0.9990
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 6.574 1      2.564
## Year              6.574 16      1.061

```

Residuals from last author



```
## [1] "List of 0 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
## -0.8893 -0.1593 0.0385 0.2222 1.8383
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.03433 0.20348 5.08 4e-06 ***
## LastAuthorFemale1 -0.34093 0.14628 -2.33 0.023 *
## Year1997 -0.18304 0.22200 -0.82 0.413
## Year1998 -0.31483 0.20530 -1.53 0.130
## Year1999 0.62217 0.88645 0.70 0.486
## Year2000 -0.00921 0.23942 -0.04 0.969
## Year2001 0.07595 0.26594 0.29 0.776
## Year2002 -0.14236 0.21658 -0.66 0.514
## Year2003 0.23556 0.24960 0.94 0.349
## Year2004 0.38883 0.25475 1.53 0.132
## Year2005 0.25691 0.36386 0.71 0.483
## Year2006 -0.01082 0.33790 -0.03 0.975
```

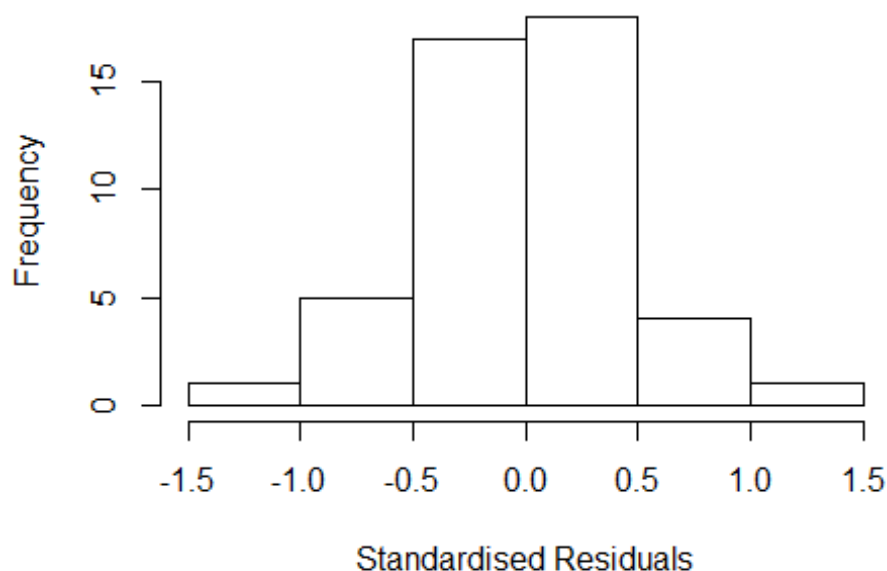
```

## Year2007      0.17402      0.30090      0.58      0.565
## Year2008     -0.02085      0.21932     -0.10      0.925
## Year2009     -0.19317      0.31696     -0.61      0.545
## Year2010     -0.05550      0.30473     -0.18      0.856
## Year2011      0.15422      0.34139      0.45      0.653
## Year2012      0.00137      0.34327      0.00      0.997
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.199, Adjusted R-squared:  -0.0315
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0502 0.8770 0.9720 0.9040 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      1.30e-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: 77"
## [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
##    3   13   4   9   9   8   1   4   6   6   13   11   4   10   13
## 2011 2012
##    10   15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    3    1    1    2    2    1    1    2    2    4    5    2    7    7
## 2011 2012

```

```
##      4      9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      0      3      1      1      2      1      0      1      2      1      4      4      2      6      6
## 2011 2012
##      4      8
## [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: 1"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.944e+01  1          NaN
## LastAuthorFemale  7.831e+15  1          8.849e+07
## UniqueAuthors    -4.295e+17  4          NaN
## Year              1.580e+32 14          1.412e+01
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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,
```

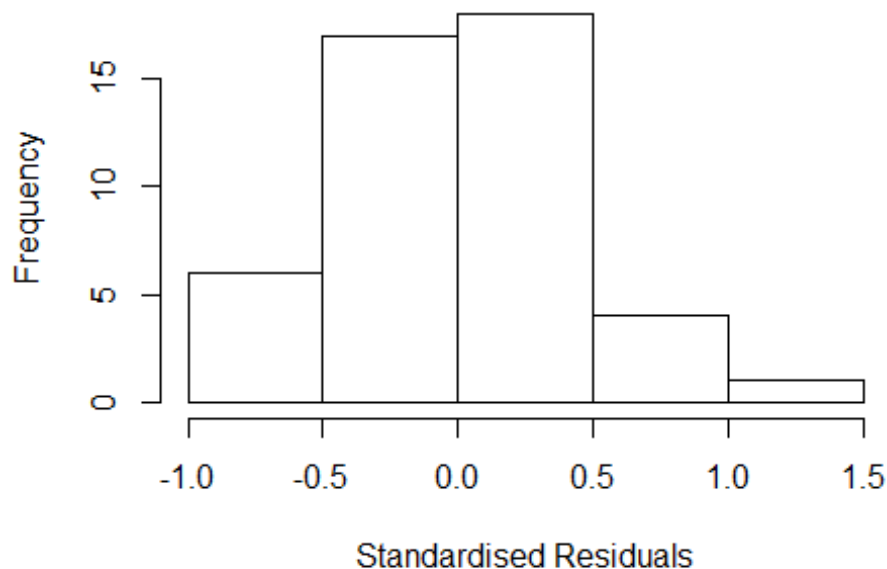
```

##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q   Median      3Q      Max
## -1.20e+00 -2.29e-01  5.82e-05  2.13e-01  1.12e+00
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.33440    0.40020   3.33  0.0027 **
## FirstAuthorFemale1  0.12804    0.16613   0.77  0.4481
## LastAuthorFemale1 -0.02702    0.22707  -0.12  0.9062
## UniqueAuthors2    -0.49884    0.25578  -1.95  0.0624 .
## UniqueAuthors3    -0.17092    0.29178  -0.59  0.5633
## UniqueAuthors4     0.00071    0.35277   0.00  0.9984
## UniqueAuthors5    -0.32329    0.46614  -0.69  0.4944
## Year1998           0.78847    0.38165   2.07  0.0493 *
## Year1999           0.98860    0.40020   2.47  0.0207 *
## Year2000          -0.52395    0.40383  -1.30  0.2063
## Year2001           0.48844    0.31165   1.57  0.1296
## Year2003           0.76944    0.31165   2.47  0.0207 *
## Year2004           0.32142    0.37531   0.86  0.3999
## Year2005           0.53347    0.38165   1.40  0.1745
## Year2006           0.05495    0.47951   0.11  0.9097
## Year2007          -0.18056    0.38950  -0.46  0.6470
## Year2008           0.46400    0.38917   1.19  0.2443
## Year2009          -0.13550    0.54497  -0.25  0.8057
## Year2010           0.00569    0.37018   0.02  0.9879
## Year2011          -0.01796    0.32839  -0.05  0.9568
## Year2012          -0.13363    0.36131  -0.37  0.7146
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.57
## Multiple R-squared:  0.381, Adjusted R-squared:  -0.114
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 36 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.637  0.924  0.976  0.943  0.988  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      2.17e-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 1.41e+02 1 1.187e+01
## LastAuthorFemale 4.35e+13 1 6.595e+06
## Year 2.58e+16 14 3.856e+00
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -0.8631 -0.2267 0.0034 0.2093 1.4519
##
## Coefficients:
```

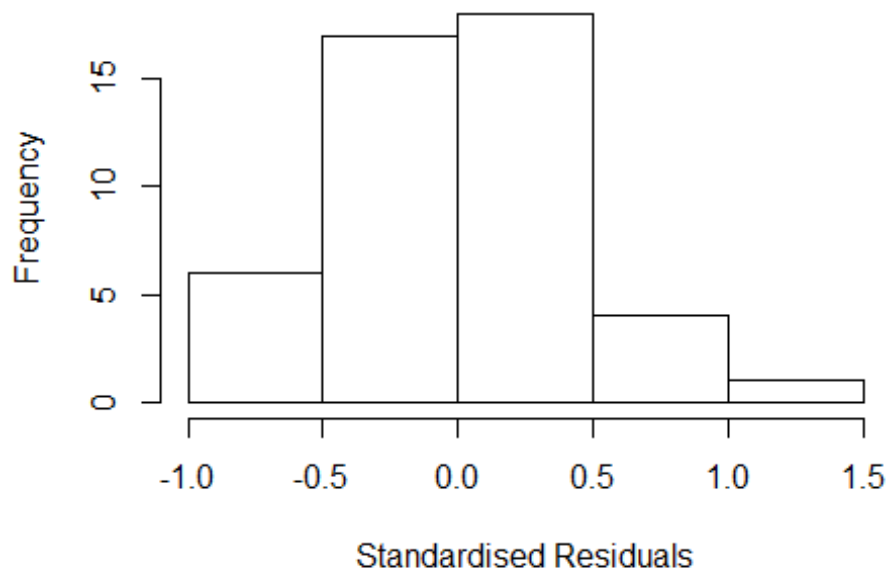


```

##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0946      0.4273   2.56  0.0159 *
## FirstAuthorFemale1 0.0636      0.1488   0.43  0.6723
## LastAuthorFemale1 0.0488      0.2174   0.22  0.8239
## Year1998          0.4536      0.4926   0.92  0.3647
## Year1999          1.2284      0.4273   2.87  0.0075 **
## Year2000          -0.3753      0.4769  -0.79  0.4377
## Year2001           0.2294      0.4273   0.54  0.5954
## Year2003           0.5104      0.4273   1.19  0.2420
## Year2004           0.0946      0.4586   0.21  0.8380
## Year2005           0.1986      0.4926   0.40  0.6898
## Year2006           0.0220      0.5238   0.04  0.9668
## Year2007          -0.2471      0.4699  -0.53  0.6030
## Year2008           0.4866      0.4340   1.12  0.2713
## Year2009          -0.2315      0.5949  -0.39  0.7000
## Year2010          -0.1769      0.4529  -0.39  0.6989
## Year2011          -0.1531      0.4553  -0.34  0.7391
## Year2012          -0.0601      0.5156  -0.12  0.9080
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.476
## Multiple R-squared:  0.315, Adjusted R-squared:  -0.0633
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.333  0.889  0.978  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      2.17e-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 224.1 1      14.969
## Year              224.1 14      1.213

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.85733 -0.22626  0.00356  0.21446  1.46875
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0881     0.4311   2.52  0.0171 *
## FirstAuthorFemale1  0.0643     0.1498   0.43  0.6710
## Year1998          0.5089     0.4311   1.18  0.2471
## Year1999          1.2349     0.4311   2.86  0.0076 **
## Year2000         -0.3447     0.4534  -0.76  0.4530
## Year2001          0.2359     0.4311   0.55  0.5882
## Year2003          0.5169     0.4311   1.20  0.2399
## Year2004          0.1008     0.4625   0.22  0.8289
## Year2005          0.2539     0.4311   0.59  0.5602
## Year2006          0.0323     0.5263   0.06  0.9515
## Year2007         -0.2399     0.4733  -0.51  0.6159
## Year2008          0.4928     0.4379   1.13  0.2693
```

```

## Year2009          -0.2418      0.6012   -0.40   0.6904
## Year2010          -0.1711      0.4569   -0.37   0.7107
## Year2011          -0.1466      0.4591   -0.32   0.7517
## Year2012          -0.0445      0.5083   -0.09   0.9308
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.463
## Multiple R-squared:  0.322, Adjusted R-squared:  -0.0171
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.293  0.883  0.977  0.909  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      2.17e-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 -6.212e+14  1      NaN
## Year             -6.212e+14 14      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
## -0.8832 -0.2343  0.0125  0.1851  1.4654
##
## Coefficients:

```

```

##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.07849    0.44200    2.44  0.0208 *
## LastAuthorFemale1 0.04318    0.20509    0.21  0.8347
## Year1998          0.47533    0.50481    0.94  0.3539
## Year1999          1.24451    0.44200    2.82  0.0085 **
## Year2000          -0.32458    0.47609   -0.68  0.5006
## Year2001          0.24551    0.44200    0.56  0.5827
## Year2003          0.52651    0.44200    1.19  0.2429
## Year2004          0.14251    0.45914    0.31  0.7584
## Year2005          0.22033    0.50481    0.44  0.6656
## Year2006          0.06769    0.54057    0.13  0.9012
## Year2007          -0.21125    0.48444   -0.44  0.6659
## Year2008          0.53451    0.44235    1.21  0.2364
## Year2009          -0.22885    0.59697   -0.38  0.7042
## Year2010          -0.14103    0.46153   -0.31  0.7620
## Year2011          -0.13696    0.46968   -0.29  0.7726
## Year2012          -0.00377    0.52905   -0.01  0.9944
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.446
## Multiple R-squared:  0.328, Adjusted R-squared:  -0.00795
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.258  0.879  0.971  0.902  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      2.17e-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: 46"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3107"
## [1] "#####"

```

```

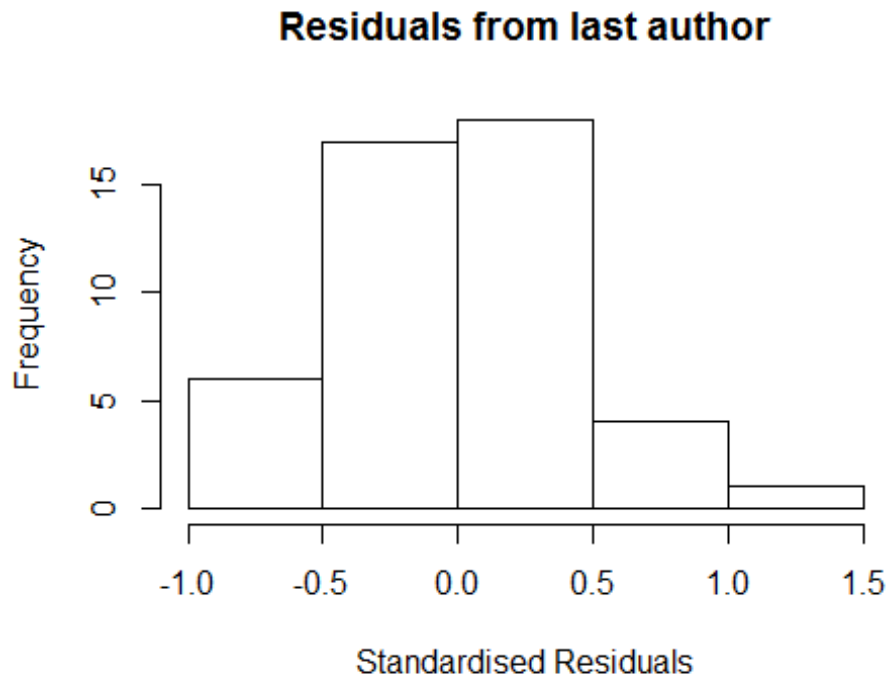
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011
## 28 26 29 8 29 32 29 19 28 1 24 20 35 44 35
## 2012
## 30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011
## 9 7 11 5 8 7 12 9 12 1 10 7 11 21 17
## 2012
## 15
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2005 2006 2007 2008 2009 2010 2011
## 8 7 9 5 7 6 12 9 11 1 8 7 10 19 14
## 2012
## 12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.93242414866094"
## [1] "Male first author team size 2018 geometric mean: 2.29855203398175"

## 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 = 52, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.93242414866094"
## [1] "Male last author team size 2018 geometric mean: 2.29855203398175"

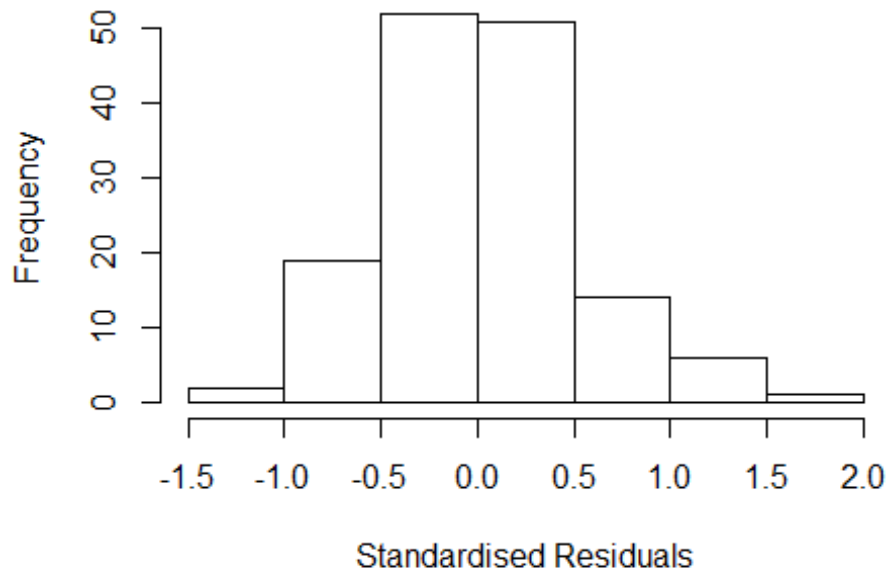
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 52, 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 2.443 1      1.563
## LastAuthorFemale 2.525 1      1.589
## UniqueAuthors    20.219 4      1.456
## Year              29.352 15     1.119
```

Residuals from first and last author and team size



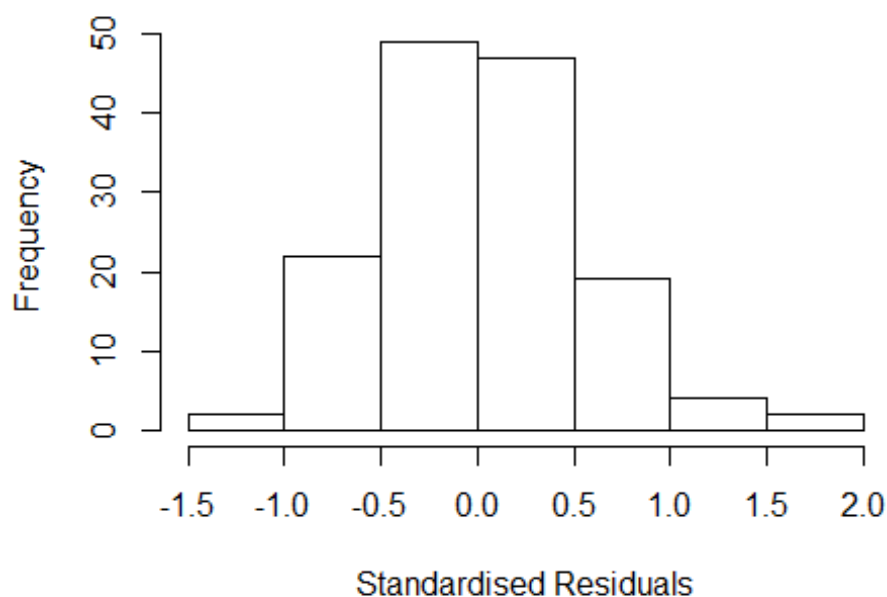
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.21e+00 -3.30e-01 8.33e-16 3.45e-01 1.59e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.97978 0.22298 4.39 2.4e-05 ***
## FirstAuthorFemale1 -0.20819 0.14131 -1.47 0.143
## LastAuthorFemale1 0.08668 0.20997 0.41 0.680
## UniqueAuthors2 0.27326 0.13444 2.03 0.044 *
## UniqueAuthors3 0.22937 0.13047 1.76 0.081 .
## UniqueAuthors4 0.19057 0.17811 1.07 0.287
## UniqueAuthors5 0.11582 0.19230 0.60 0.548
## Year1997 0.10684 0.38680 0.28 0.783
## Year1998 0.22386 0.31088 0.72 0.473
## Year1999 0.37746 0.26576 1.42 0.158
```

```

## Year2000      0.20485    0.38062    0.54    0.591
## Year2001      0.18750    0.29611    0.63    0.528
## Year2002     -0.12215    0.27640   -0.44    0.659
## Year2003     -0.23697    0.24722   -0.96    0.340
## Year2005     -0.00631    0.27729   -0.02    0.982
## Year2006     -0.44915    0.21745   -2.07    0.041 *
## Year2007     -0.09141    0.23594   -0.39    0.699
## Year2008     -0.05174    0.23455   -0.22    0.826
## Year2009     -0.28746    0.25055   -1.15    0.253
## Year2010      0.00337    0.23615    0.01    0.989
## Year2011     -0.12655    0.24909   -0.51    0.612
## Year2012     -0.05845    0.26913   -0.22    0.828
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.505
## Multiple R-squared:  0.149, Adjusted R-squared:  0.00315
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 126 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.299  0.875   0.946   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      6.90e-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 2.199 1 1.483
## LastAuthorFemale 1.856 1 1.362
## Year 2.456 15 1.030

```


Residuals from first and last author



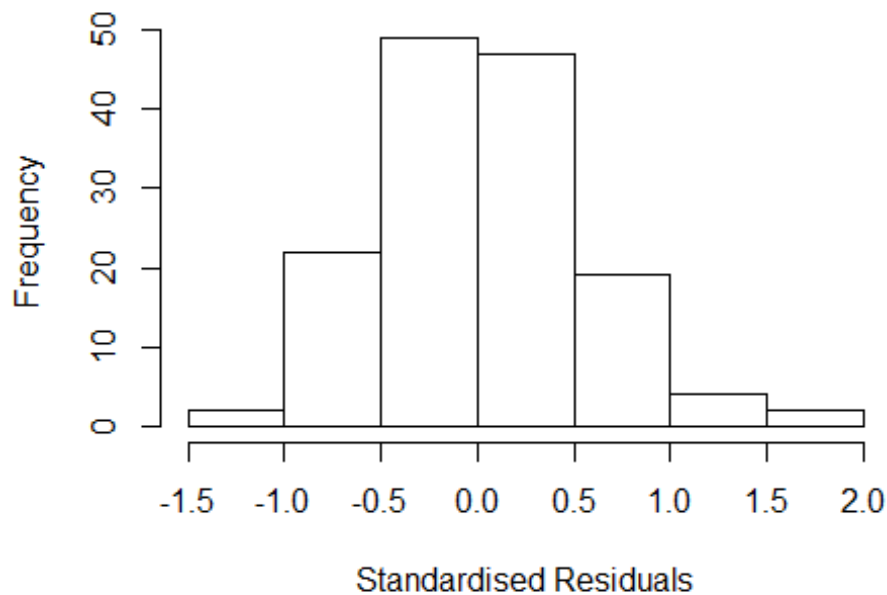
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.08e+00 -3.42e-01 -2.78e-16 3.17e-01 1.60e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.15223 0.21033 5.48 2.2e-07 ***
## FirstAuthorFemale1 -0.20816 0.14333 -1.45 0.149
## LastAuthorFemale1 0.05166 0.20263 0.25 0.799
## Year1997 0.19765 0.37390 0.53 0.598
## Year1998 0.23661 0.30911 0.77 0.445
## Year1999 0.43991 0.24681 1.78 0.077 .
## Year2000 0.21711 0.37616 0.58 0.565
## Year2001 0.24544 0.31040 0.79 0.431
## Year2002 -0.06180 0.25363 -0.24 0.808
## Year2003 -0.24644 0.24028 -1.03 0.307
## Year2005 -0.04838 0.29987 -0.16 0.872
## Year2006 -0.39223 0.21033 -1.86 0.065 .
```

```

## Year2007          -0.11048    0.24123   -0.46    0.648
## Year2008          -0.07173    0.23401   -0.31    0.760
## Year2009          -0.28284    0.27203   -1.04    0.300
## Year2010          -0.00601    0.24306   -0.02    0.980
## Year2011          -0.12076    0.25083   -0.48    0.631
## Year2012          -0.06048    0.28365   -0.21    0.831
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.108, Adjusted R-squared:  -0.0113
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.325  0.876  0.959  0.912  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      6.90e-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 1.823 1          1.35
## Year              1.823 15          1.02

```

Residuals from first author



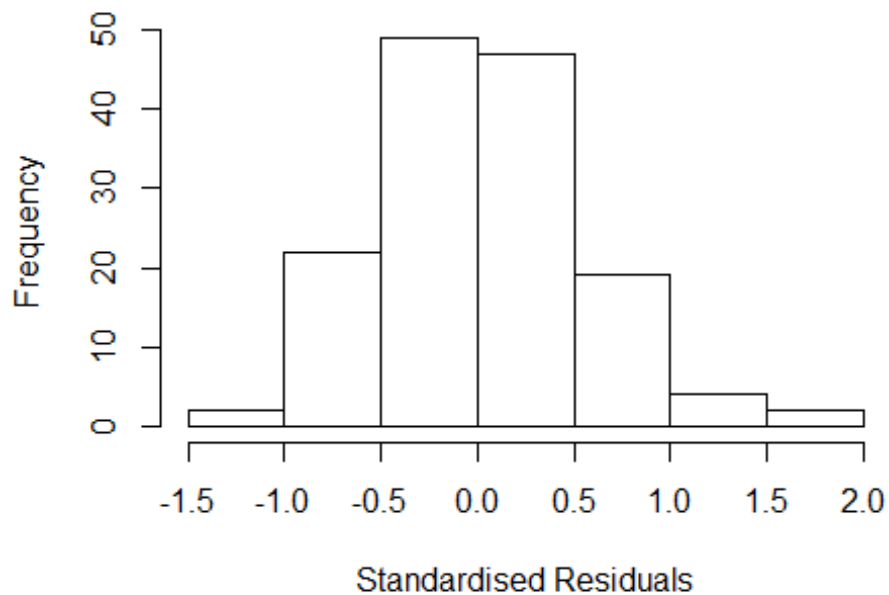
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.06e+00 -3.48e-01 3.89e-16 3.05e-01 1.60e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.151057 0.209452 5.50 2e-07 ***
## FirstAuthorFemale1 -0.195795 0.129013 -1.52 0.132
## Year1997 0.204296 0.373712 0.55 0.586
## Year1998 0.237443 0.311512 0.76 0.447
## Year1999 0.436188 0.244444 1.78 0.077 .
## Year2000 0.214383 0.373223 0.57 0.567
## Year2001 0.254118 0.314976 0.81 0.421
## Year2002 -0.060645 0.252922 -0.24 0.811
## Year2003 -0.245398 0.239533 -1.02 0.308
## Year2005 -0.043774 0.297827 -0.15 0.883
## Year2006 -0.391057 0.209452 -1.87 0.064 .
## Year2007 -0.103963 0.238939 -0.44 0.664
```

```

## Year2008          -0.072257    0.232654   -0.31    0.757
## Year2009          -0.275561    0.273775   -1.01    0.316
## Year2010           0.000763    0.239946    0.00    0.997
## Year2011          -0.115890    0.250609   -0.46    0.645
## Year2012          -0.053972    0.283713   -0.19    0.849
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.107, Adjusted R-squared:  -0.0043
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 135 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.325  0.874  0.956  0.910  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      6.90e-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 1.445 1          1.202
## Year            1.445 15          1.012

```

Residuals from last author



```
## [1] "List of 0 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
## -1.05501 -0.33060 -0.00164 0.32643 1.69279
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1309 0.1933 5.85 3.9e-08 ***
## LastAuthorFemale1 -0.0169 0.2163 -0.08 0.938
## Year1997 0.1283 0.3936 0.33 0.745
## Year1998 0.2571 0.3027 0.85 0.397
## Year1999 0.3786 0.2232 1.70 0.092 .
## Year2000 0.1664 0.3282 0.51 0.613
## Year2001 0.2766 0.3064 0.90 0.368
## Year2002 -0.0406 0.2396 -0.17 0.866
## Year2003 -0.2255 0.2259 -1.00 0.320
## Year2005 -0.0758 0.2782 -0.27 0.786
## Year2006 -0.3709 0.1933 -1.92 0.057 .
## Year2007 -0.1084 0.2273 -0.48 0.634
```

```

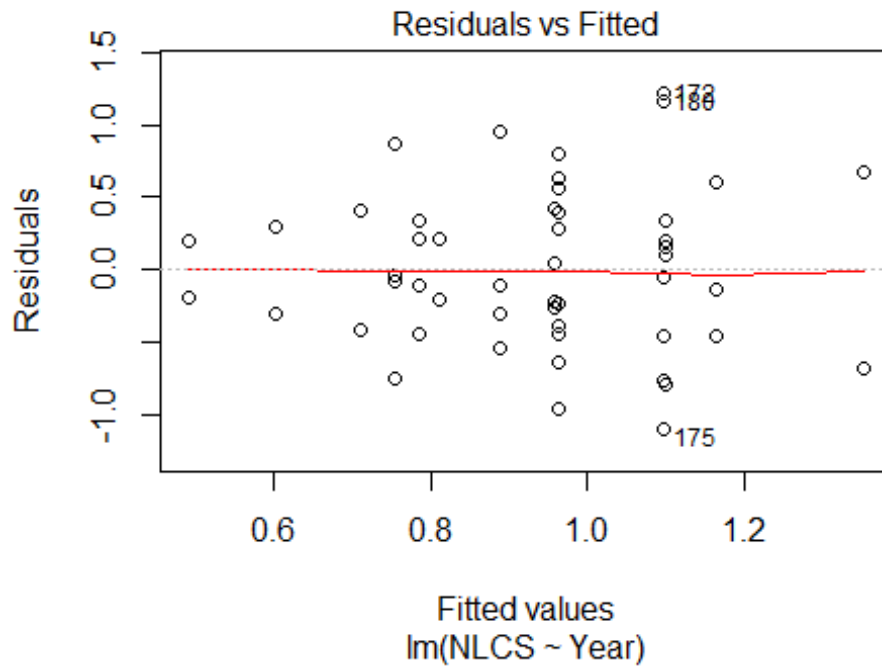
## Year2008          -0.0804      0.2120    -0.38      0.705
## Year2009          -0.2856      0.2737    -1.04      0.299
## Year2010           0.0107      0.2295      0.05      0.963
## Year2011          -0.1115      0.2373    -0.47      0.639
## Year2012          -0.0913      0.2687    -0.34      0.735
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.517
## Multiple R-squared:  0.0943, Adjusted R-squared:  -0.0189
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 137 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.262  0.872  0.954  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      6.90e-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: 145"
## [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 2004 2005 2006 2007 2009 2010 2011 2012
##    3    2    3    4    2    3    1    3    1    2    4    2    1    2    3
##
## 1996 1997 1998 1999 2000 2001 2002 2004 2005 2006 2007 2009 2010 2011 2012
##    1    1    3    2    1    0    1    2    1    1    2    2    1    1    1
##
## 1996 1997 1998 1999 2000 2001 2002 2004 2005 2006 2007 2009 2010 2011 2012
##    1    1    2    2    1    0    1    2    1    1    2    1    1    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with

```

```

these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [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: 18"
## [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
##    3    2    4    4    5   11    6    5   14    9    9   16    9    9    3
## 2011 2012
##    4    6
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    0    2    0    0    4    2    4    6    4    2    5    4    6    2
## 2011 2012
##    3    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    0    2    0    0    4    2    2    6    4    2    4    4    4    2
## 2011 2012
##    2    4
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 7.9, df = 13, p-value = 0.8

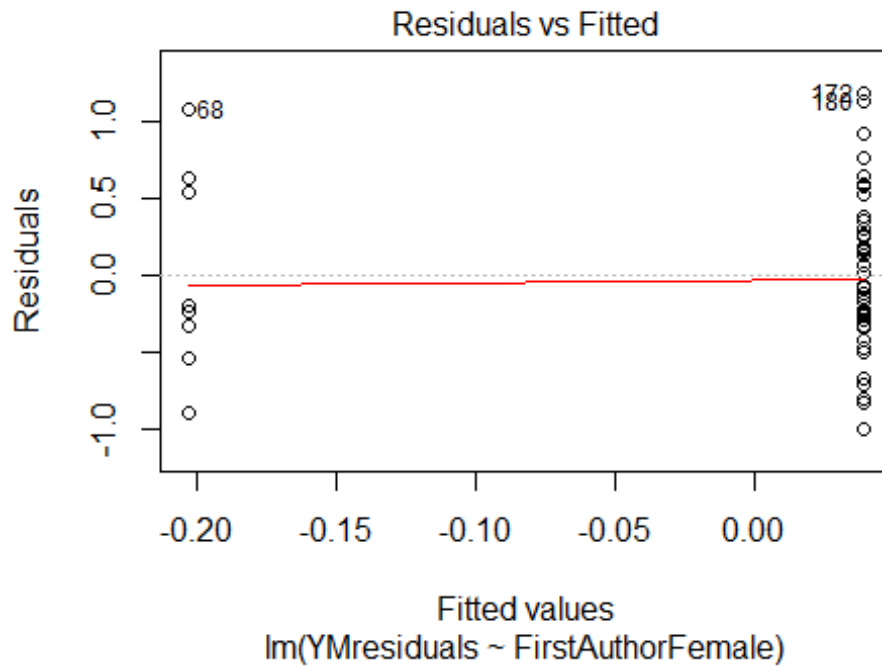
```



```
##
## 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: 5"
## [1] "Male first author team size 2018 geometric mean: 1.68179283050743"

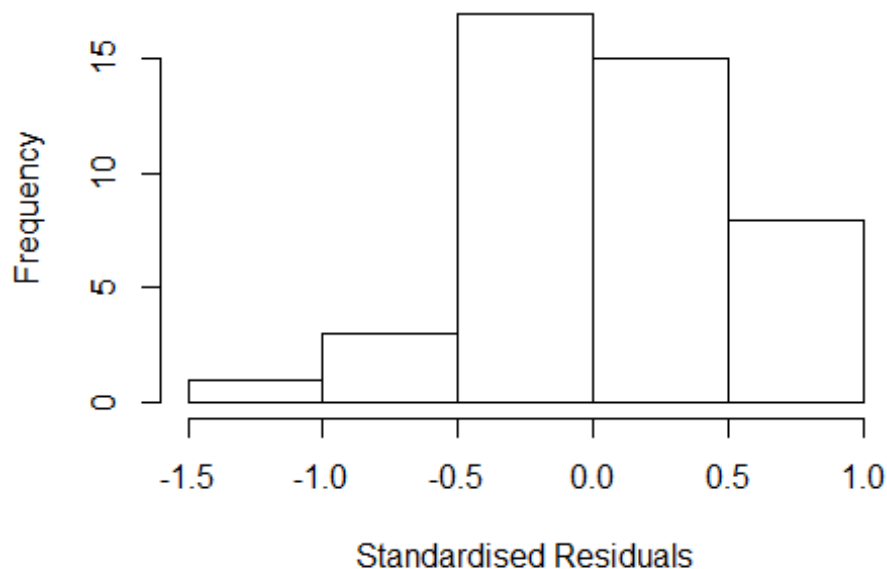
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

```
##
## 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] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 2.09127910518255"
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
##
```

	GVIF	Df	GVIF^(1/(2*Df))
FirstAuthorFemale	3.096	1	1.760
LastAuthorFemale	1.903	1	1.380
UniqueAuthors	24.474	4	1.491
Year	38.936	13	1.151

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.2172 -0.2758 0.0595 0.2593 0.7728
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2639 0.1600 7.90 3.9e-08 ***
## FirstAuthorFemale1 -0.7709 0.1838 -4.19 0.00032 ***
## LastAuthorFemale1 0.6877 0.4995 1.38 0.18130
## UniqueAuthors2 -0.3359 0.1559 -2.15 0.04142 *
## UniqueAuthors3 0.0357 0.1664 0.21 0.83205
## UniqueAuthors4 -0.0191 0.3952 -0.05 0.96176
## UniqueAuthors5 0.0452 0.5057 0.09 0.92956
## Year1998 -0.1170 0.2145 -0.55 0.59047
## Year2001 -0.3652 0.1822 -2.00 0.05643 .
## Year2002 -0.6055 0.3060 -1.98 0.05943 .
```

```

## Year2003          -0.3975      0.1935    -2.05  0.05104 .
## Year2004          -0.1965      0.3011    -0.65  0.52038
## Year2005          -0.3053      0.2708    -1.13  0.27075
## Year2006          -0.6609      0.2737    -2.41  0.02374 *
## Year2007          -0.1676      0.2241    -0.75  0.46196
## Year2008          -0.2665      0.4293    -0.62  0.54059
## Year2009           0.2803      0.6453      0.43  0.66786
## Year2010           0.0886      0.6349      0.14  0.89017
## Year2011           0.3015      0.2028      1.49  0.15000
## Year2012          -0.0338      0.2201    -0.15  0.87933
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.632
## Multiple R-squared:  0.307, Adjusted R-squared:  -0.242
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
##  one weight is ~= 1. The remaining 43 ones are summarized as
##    Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.691  0.939  0.979  0.952  0.991  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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 44.900 1      6.701
## LastAuthorFemale  1.449 1      1.204
## Year              47.808 13      1.160

## [1] "List of 0 outliers with 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 =

```

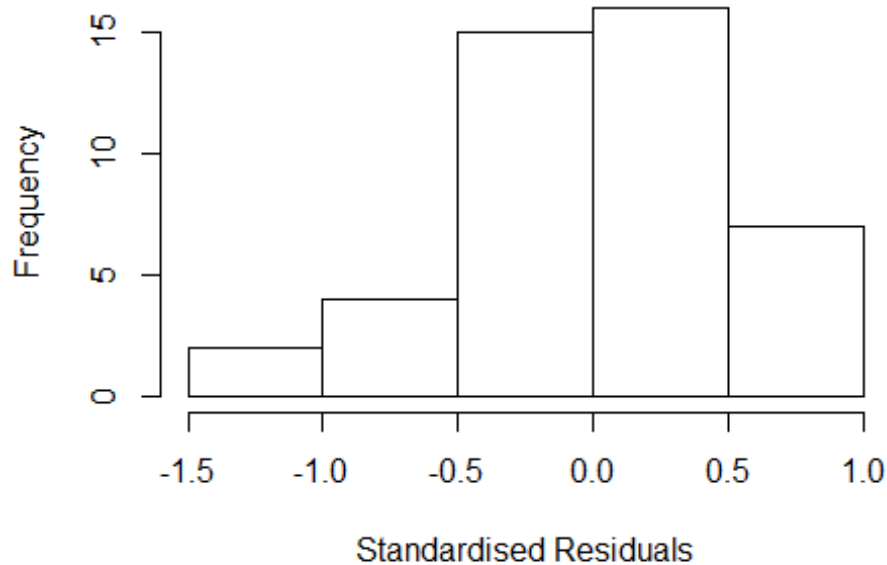
```

AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min        1Q    Median        3Q        Max
## -1.2320 -0.3196  0.0322  0.3088  0.8078
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0802     0.1142   9.46 3.3e-10 ***
## FirstAuthorFemale1 -0.7394     0.2206  -3.35  0.0023 **
## LastAuthorFemale1  0.7237     0.4161   1.74  0.0930 .
## Year1998          -0.2692     0.1878  -1.43  0.1627
## Year2001          -0.2892     0.1885  -1.53  0.1362
## Year2002          -0.5897     0.1815  -3.25  0.0030 **
## Year2003          -0.3817     0.1153  -3.31  0.0026 **
## Year2004          -0.1260     0.2647  -0.48  0.6376
## Year2005          -0.1227     0.1715  -0.72  0.4801
## Year2006          -0.4772     0.2511  -1.90  0.0678 .
## Year2007          -0.0402     0.2523  -0.16  0.8746
## Year2008          -0.0757     0.4304  -0.18  0.8616
## Year2009           0.4788     0.6933   0.69  0.4955
## Year2010           0.2723     0.6565   0.41  0.6815
## Year2011           0.3173     0.3092   1.03  0.3136
## Year2012          -0.0296     0.2206  -0.13  0.8942
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.264, Adjusted R-squared:  -0.13
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 38 ones are summarized as
##      Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##      0.639  0.901  0.963  0.933  0.982  0.996
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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"
## 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
```

Residuals from first and last author



```
## [1] "Regression 4: Last author gender, Year as factors"
##               GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.07  1      1.034
## Year             1.07 13      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
## -1.22e+00 -3.77e-01 -1.28e-16  3.43e-01  1.00e+00
##
## Coefficients:
```

```

##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.7105     0.3165   2.24   0.033 *
## LastAuthorFemale1 -0.0158     0.3526  -0.04   0.965
## Year1998          0.1005     0.3497   0.29   0.776
## Year2001          0.0802     0.3493   0.23   0.820
## Year2002         -0.2200     0.3464  -0.64   0.530
## Year2003         -0.0120     0.3169  -0.04   0.970
## Year2004          0.2449     0.3964   0.62   0.542
## Year2005          0.2474     0.3421   0.72   0.475
## Year2006         -0.1075     0.3868  -0.28   0.783
## Year2007          0.3266     0.3834   0.85   0.401
## Year2008          0.2880     0.5046   0.57   0.573
## Year2009          0.8359     0.6974   1.20   0.240
## Year2010          0.6420     0.6943   0.92   0.363
## Year2011          0.6870     0.4256   1.61   0.117
## Year2012          0.1313     0.4427   0.30   0.769
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.624
## Multiple R-squared:  0.212, Adjusted R-squared:  -0.169
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 4 weights are ~= 1. The remaining 40 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.683  0.916  0.964  0.940  0.984  0.997
## Algorithmic parameters:
##           tuning.chi           bb           tuning.psi           refine.tol
##           1.55e+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-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: 44"
## [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
##    1    6    5    7    4    2    9    1   17   14   15   13   27   15   10
## 2011 2012
##    6    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    3    4    3    2    1    5    0    6    8    7    8   12    8    8
## 2011 2012
##    2    2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    2    1    2    1    3    0    4    7    4    5    8    6    4
## 2011 2012
##    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male first author team size 2018 geometric mean: 4.94923200383977"

## 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 = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 4.74288121955862"

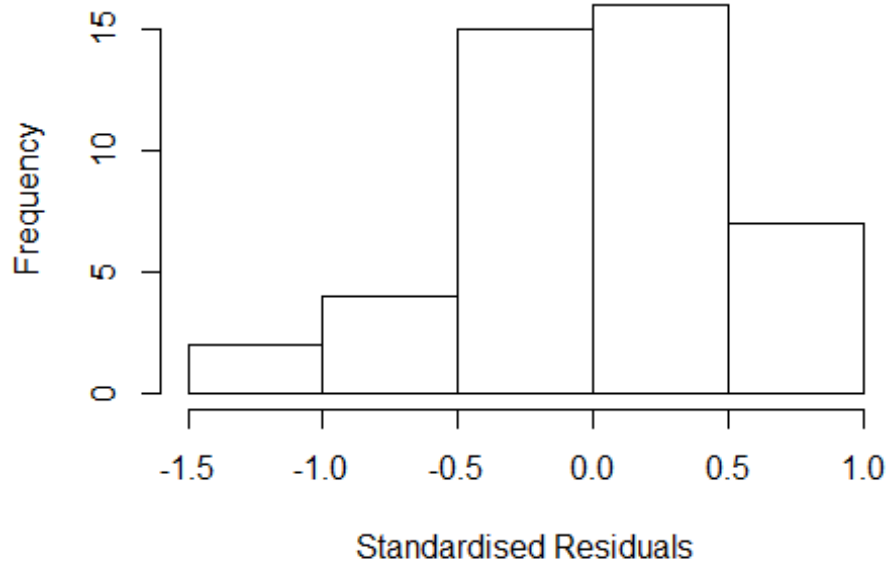
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, 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"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

```

Residuals from last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale -1.012e+16  1      NaN
## LastAuthorFemale  -7.410e+15  1      NaN
## UniqueAuthors     1.589e+45  4    446833.9
## Year               2.134e+73 15     278.2

## [1] "List of 0 outliers with 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
## -7.28e-01 -1.74e-01  3.89e-16  2.01e-01  7.89e-01
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.2890     0.0000      Inf <2e-16 ***
## FirstAuthorFemale1  0.1578     0.1590     0.99  0.329
## LastAuthorFemale1  0.4754     0.2785     1.71  0.098 .
## UniqueAuthors2    0.0118     0.2741     0.04  0.966
```



```

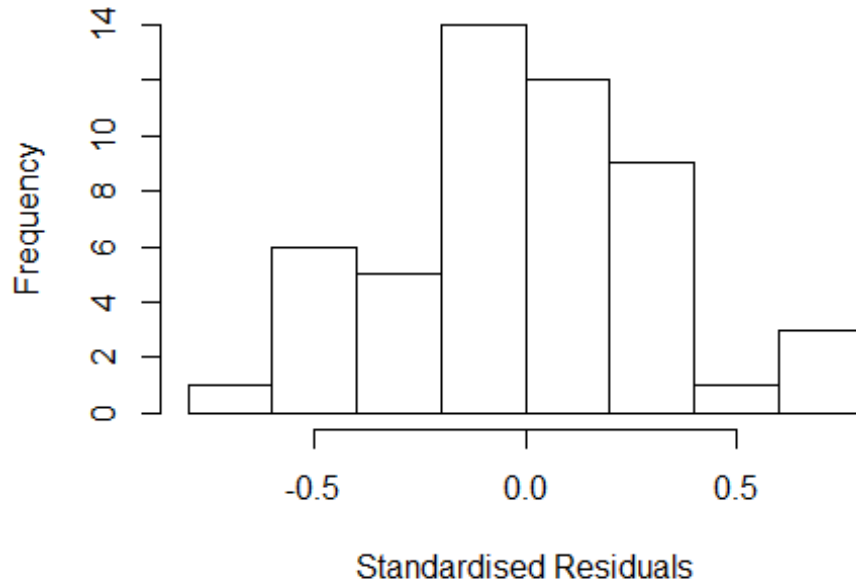
## UniqueAuthors3      0.2350      0.2533      0.93      0.361
## UniqueAuthors4      0.0536      0.3281      0.16      0.871
## UniqueAuthors5      0.0555      0.2971      0.19      0.853
## Year1997             0.3904      0.3172      1.23      0.228
## Year1998            -0.4554      0.4050     -1.12      0.270
## Year1999            -0.4760      0.2533     -1.88      0.070 .
## Year2000            -0.5882      0.4560     -1.29      0.207
## Year2001            -0.1800      0.3918     -0.46      0.649
## Year2002            -0.0956      0.2729     -0.35      0.729
## Year2004            -0.0321      0.2877     -0.11      0.912
## Year2005            -0.4888      0.3104     -1.57      0.126
## Year2006            -0.6808      0.2533     -2.69      0.012 *
## Year2007            -0.4530      0.3038     -1.49      0.147
## Year2008            -0.3501      0.2092     -1.67      0.105
## Year2009            -0.5520      0.2494     -2.21      0.035 *
## Year2010            -0.9602      0.5975     -1.61      0.119
## Year2011            -0.3398      0.2741     -1.24      0.225
## Year2012            -0.6094      0.3389     -1.80      0.083 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.419
## Multiple R-squared:  0.429, Adjusted R-squared:  0.016
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 41 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.704  0.913  0.965  0.935  0.990  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      1.96e-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

```

```
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author and team size



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## 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 ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.79446 -0.20505  0.00778  0.20029  0.66585
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.29e+00   1.77e-08  7.30e+07 < 2e-16 ***
## FirstAuthorFemale1 1.50e-01   1.44e-01  1.04e+00  0.30515
## LastAuthorFemale1  4.27e-01   2.01e-01  2.13e+00  0.04105 *
```

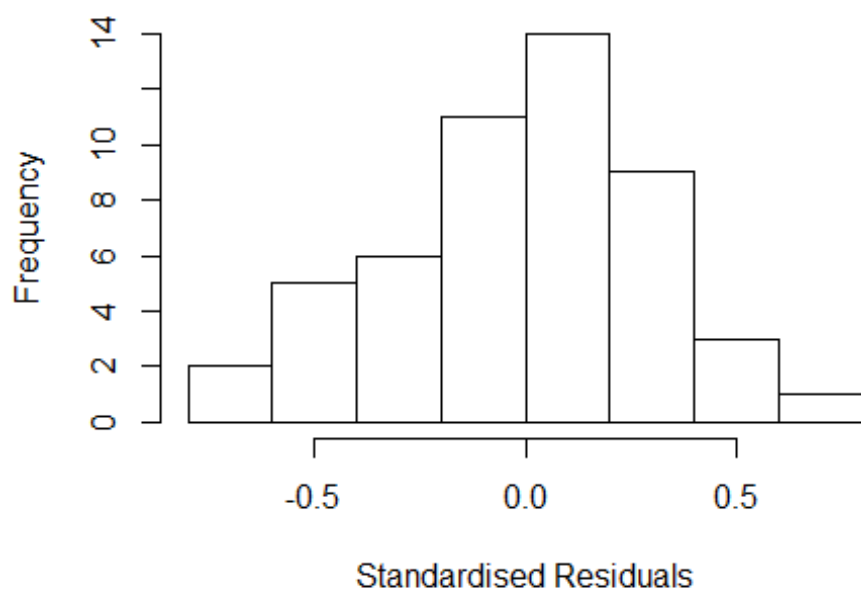
```

## Year1997          4.10e-01    1.44e-01    2.84e+00    0.00772 **
## Year1998          -2.17e-01    3.21e-01   -6.80e-01    0.50415
## Year1999          -2.41e-01    0.00e+00      -Inf    < 2e-16 ***
## Year2000          -5.30e-01    3.99e-01   -1.33e+00    0.19303
## Year2001          -1.12e-01    2.80e-01   -4.00e-01    0.69059
## Year2002           1.42e-01    1.01e-01    1.41e+00    0.16892
## Year2004           1.76e-02    1.92e-01    9.00e-02    0.92753
## Year2005          -3.16e-01    2.33e-01   -1.36e+00    0.18447
## Year2006          -5.43e-01    7.55e-02   -7.19e+00    3.1e-08 ***
## Year2007          -3.95e-01    2.34e-01   -1.69e+00    0.10038
## Year2008          -2.84e-01    1.69e-01   -1.68e+00    0.10165
## Year2009          -4.28e-01    1.19e-01   -3.61e+00    0.00102 **
## Year2010          -7.84e-01    4.42e-01   -1.77e+00    0.08541 .
## Year2011          -3.28e-01    0.00e+00      -Inf    < 2e-16 ***
## Year2012          -5.48e-01    1.44e-01   -3.80e+00    0.00059 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.399
## Multiple R-squared:  0.407, Adjusted R-squared:  0.102
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 37 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.672  0.864  0.938  0.918  0.982  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           1.96e-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"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

```

Residuals from first and last author



```
##                               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.017e+16  1      1.008e+08
## Year              1.017e+16 15      3.416e+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 + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -7.25e-01 -1.77e-01 -3.89e-16  2.30e-01  7.08e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)   1.29e+00   0.00e+00    Inf < 2e-16 ***
## FirstAuthorFemale1 1.46e-01   1.55e-01   9.40e-01  0.3522
## Year1997        4.14e-01   1.55e-01   2.68e+00  0.0114 *
## Year1998       -2.14e-01   3.08e-01  -7.00e-01  0.4904
## Year1999       -2.41e-01   2.30e-08  -1.05e+07 < 2e-16 ***
## Year2000       -1.01e-01   3.35e-01  -3.00e-01  0.7651
## Year2001        3.19e-01   1.55e-01   2.06e+00  0.0469 *
## Year2002        1.43e-01   1.03e-01   1.39e+00  0.1735
```

```

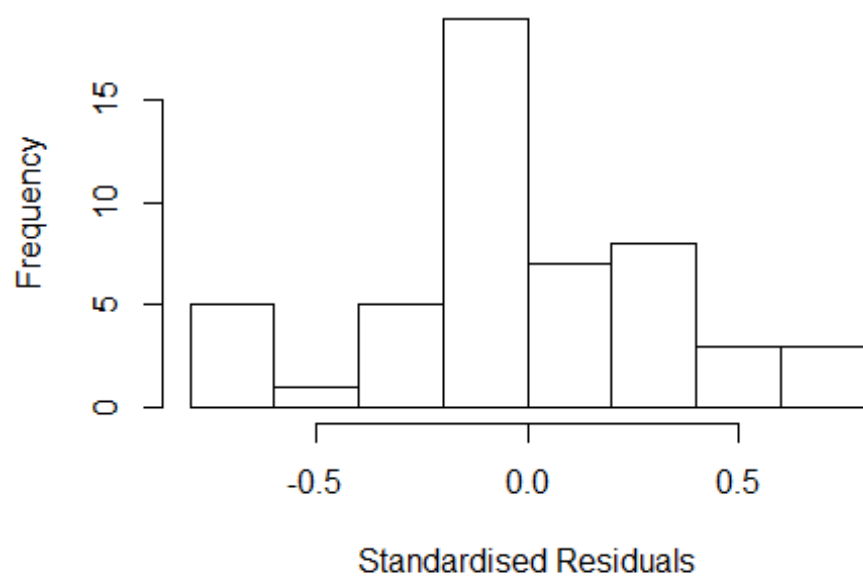
## Year2004          1.13e-01    2.48e-01    4.50e-01    0.6526
## Year2005          -2.00e-01    2.30e-01   -8.70e-01    0.3887
## Year2006          -5.40e-01    8.02e-02   -6.74e+00    9.6e-08 ***
## Year2007          -1.31e-01    2.38e-01   -5.50e-01    0.5870
## Year2008          -2.30e-01    1.91e-01   -1.20e+00    0.2366
## Year2009          -3.78e-01    1.64e-01   -2.30e+00    0.0275 *
## Year2010          -4.59e-01    2.48e-01   -1.85e+00    0.0724 .
## Year2011          -3.28e-01    1.38e-08   -2.37e+07    < 2e-16 ***
## Year2012          -5.44e-01    1.55e-01   -3.52e+00    0.0013 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.436
## Multiple R-squared:  0.303, Adjusted R-squared:  -0.0254
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
##  15 weights are ~ 1. The remaining 36 ones are summarized as
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.764  0.838  0.942   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      1.96e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

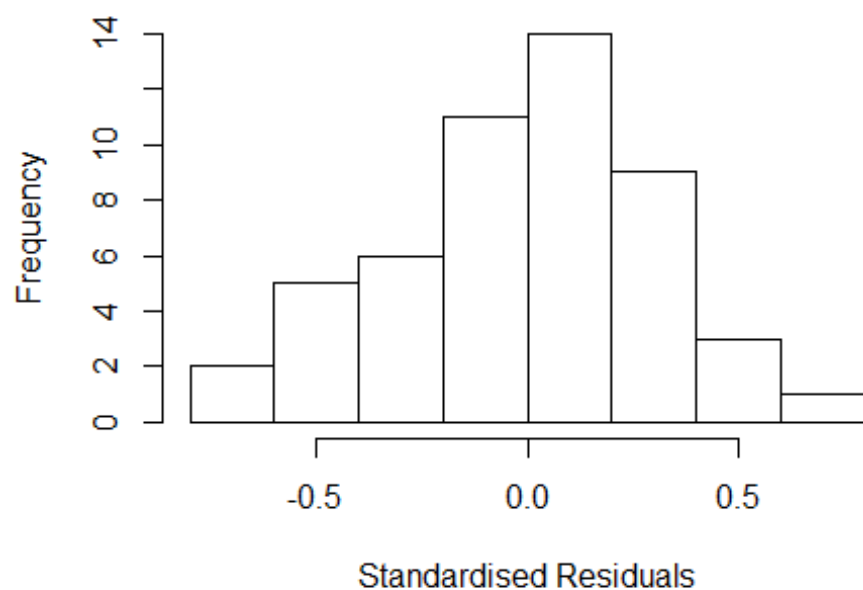
```

Residuals from first author



##	GVIF	Df	$GVIF^{1/(2 \cdot Df)}$
## LastAuthorFemale	NaN	1	NaN
## Year	NaN	15	NaN

Residuals from last author

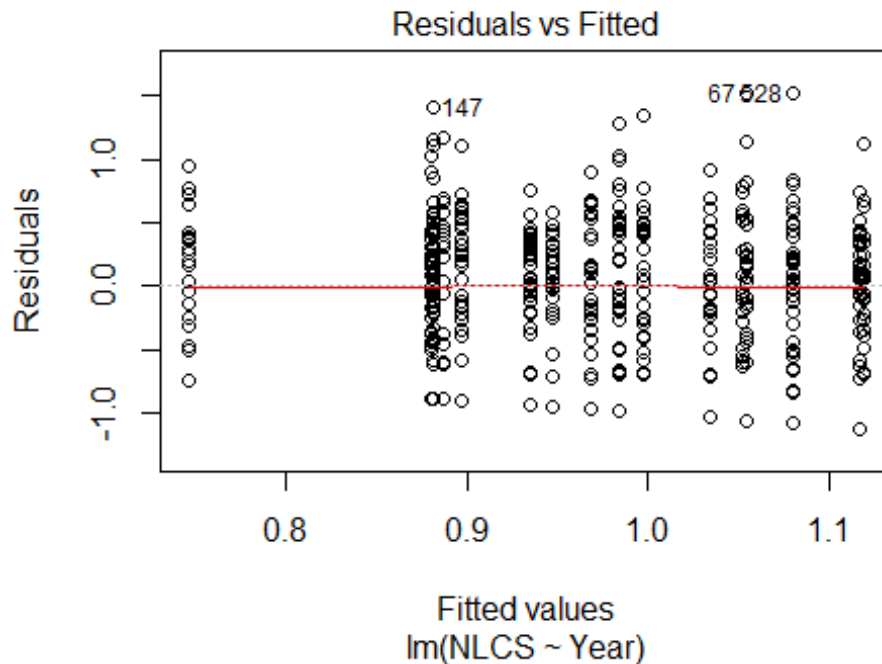


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.717 -0.185 0.000 0.237 0.593
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2890 0.0000 Inf <2e-16 ***
## LastAuthorFemale1 0.4264 0.2027 2.10 0.0429 *
## Year1997 0.5600 0.0000 Inf <2e-16 ***
## Year1998 -0.1415 0.2336 -0.61 0.5487
## Year1999 -0.2410 0.0000 -Inf <2e-16 ***
## Year2000 -0.4544 0.5062 -0.90 0.3758
## Year2001 0.0386 0.2027 0.19 0.8499
## Year2002 0.1901 0.1309 1.45 0.1555
## Year2004 0.1020 0.2049 0.50 0.6218
## Year2005 -0.1704 0.1941 -0.88 0.3863
## Year2006 -0.4675 0.0218 -21.45 <2e-16 ***
## Year2007 -0.3611 0.2531 -1.43 0.1628
## Year2008 -0.2717 0.1619 -1.68 0.1024
## Year2009 -0.3970 0.1150 -3.45 0.0015 **
## Year2010 -0.7108 0.3701 -1.92 0.0632 .
## Year2011 -0.3280 0.0000 -Inf <2e-16 ***
## Year2012 -0.3980 0.0000 -Inf <2e-16 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.374
## Multiple R-squared: 0.393, Adjusted R-squared: 0.107
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 38 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.694 0.865 0.937 0.907 0.982 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 1.96e-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: 51"
## [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
##    46   31   22   36   33   30   29   33   34   37   47   47   40   38   47
## 2011 2012
##    49   31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    38   26   17   29   25   24   27   26   29   30   42   40   33   34   38
## 2011 2012
##    41   25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    35   23   15   27   22   23   25   23   27   27   35   39   29   30   37
## 2011 2012
##    36   23
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
##      Bartlett test of homogeneity of variances
##
##      data:  NLCS by Year
##      Bartlett's K-squared = 16, df = 16, p-value = 0.4

```

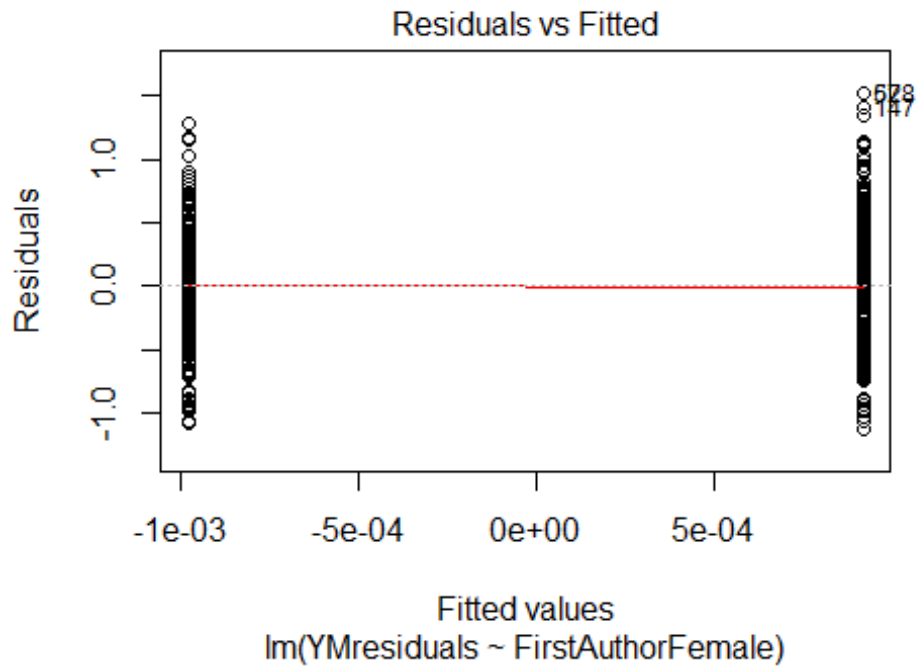
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.4, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 3.11660431257537"
## [1] "Male first author team size 2018 geometric mean: 2.12598075857908"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

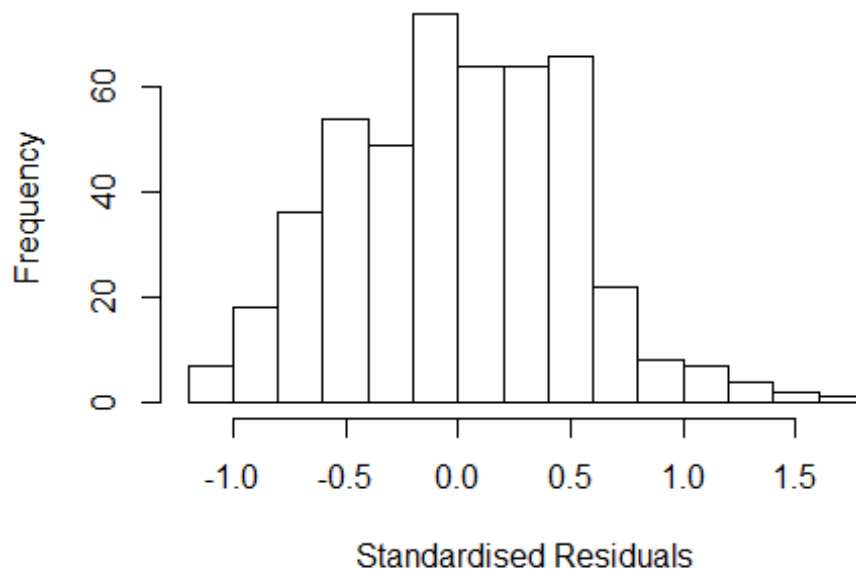
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.17659518715315"
## [1] "Male last author team size 2018 geometric mean: 2.28584117097325"

## 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 = 94, 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 1.289 1          1.136
## LastAuthorFemale  1.166 1          1.080
## UniqueAuthors    1.600 4          1.060
## Year              1.886 16         1.020
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.079324 -0.380919 -0.000379 0.369884 1.703206
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8249 0.1230 6.70 6.0e-11 ***
## FirstAuthorFemale1 -0.0342 0.0514 -0.67 0.50615
## LastAuthorFemale1 -0.0422 0.0496 -0.85 0.39581
## UniqueAuthors2 0.2708 0.0647 4.19 3.4e-05 ***
## UniqueAuthors3 0.3154 0.0714 4.42 1.3e-05 ***
## UniqueAuthors4 0.3354 0.0907 3.70 0.00025 ***
## UniqueAuthors5 0.2705 0.0892 3.03 0.00257 **
## Year1997 0.0114 0.1679 0.07 0.94578
## Year1998 0.1662 0.1849 0.90 0.36930
## Year1999 -0.1091 0.1500 -0.73 0.46748
```

```

## Year2000          0.0137      0.1381      0.10  0.92115
## Year2001         -0.0183      0.1763     -0.10  0.91729
## Year2002         -0.1241      0.1520     -0.82  0.41471
## Year2003         -0.2149      0.1650     -1.30  0.19337
## Year2004          0.2073      0.1458      1.42  0.15583
## Year2005         -0.0726      0.1346     -0.54  0.58989
## Year2006         -0.0901      0.1342     -0.67  0.50254
## Year2007          0.0689      0.1413      0.49  0.62609
## Year2008         -0.0163      0.1570     -0.10  0.91720
## Year2009         -0.1196      0.1556     -0.77  0.44242
## Year2010         -0.0106      0.1436     -0.07  0.94112
## Year2011         -0.1360      0.1387     -0.98  0.32754
## Year2012          0.1283      0.1403      0.91  0.36110
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.525
## Multiple R-squared:  0.0936, Adjusted R-squared:  0.0496
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 49 weights are ~= 1. The remaining 427 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.270  0.886  0.943  0.913  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      2.10e-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 1.157 1          1.075
## LastAuthorFemale  1.120 1          1.058
## 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)

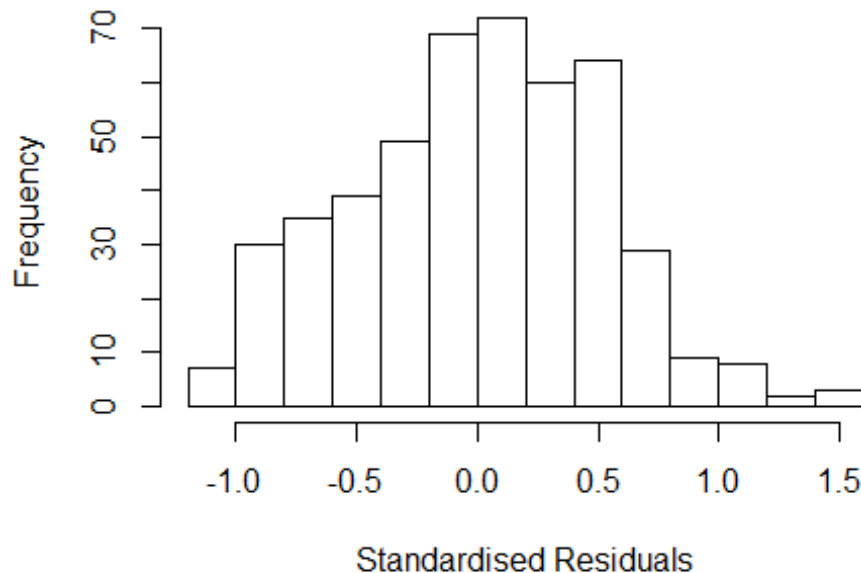
```

```
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1389 -0.3673  0.0155  0.3878  1.5358
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9829     0.1271   7.73 6.9e-14 ***
## FirstAuthorFemale1  0.0281     0.0509   0.55  0.58
## LastAuthorFemale1 -0.0423     0.0502  -0.84  0.40
## Year1997          0.0473     0.1793   0.26  0.79
## Year1998          0.1271     0.1868   0.68  0.50
## Year1999         -0.1231     0.1623  -0.76  0.45
## Year2000          0.0616     0.1464   0.42  0.67
## Year2001         -0.0935     0.1929  -0.48  0.63
## Year2002         -0.0949     0.1637  -0.58  0.56
## Year2003         -0.2517     0.1805  -1.39  0.16
## Year2004          0.1560     0.1534   1.02  0.31
## Year2005         -0.0523     0.1451  -0.36  0.72
## Year2006         -0.0692     0.1458  -0.47  0.64
## Year2007          0.1104     0.1497   0.74  0.46
## Year2008         -0.0244     0.1705  -0.14  0.89
## Year2009         -0.0795     0.1662  -0.48  0.63
## Year2010         -0.0164     0.1509  -0.11  0.91
## Year2011         -0.0668     0.1538  -0.43  0.66
## Year2012          0.1236     0.1530   0.81  0.42
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.538
## Multiple R-squared:  0.035, Adjusted R-squared:  -0.00299
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 41 weights are ~= 1. The remaining 435 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.395  0.878  0.946  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      2.10e-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"

## 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
```

Residuals from first and last author



```
## [1] "Regression 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 1.108 1           1.053
## 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 ~ LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
```

```

## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.1488 -0.3680  0.0249  0.3849  1.5266
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.9926    0.1254   7.92 1.9e-14 ***
## LastAuthorFemale1 -0.0342    0.0508  -0.67   0.50
## Year1997        0.0469    0.1800   0.26   0.79
## Year1998        0.1218    0.1853   0.66   0.51
## Year1999       -0.1223    0.1620  -0.75   0.45
## Year2000        0.0590    0.1466   0.40   0.69
## Year2001       -0.0954    0.1919  -0.50   0.62
## Year2002       -0.0952    0.1640  -0.58   0.56
## Year2003       -0.2556    0.1807  -1.41   0.16
## Year2004        0.1562    0.1532   1.02   0.31
## Year2005       -0.0483    0.1447  -0.33   0.74
## Year2006       -0.0670    0.1456  -0.46   0.65
## Year2007        0.1112    0.1495   0.74   0.46
## Year2008       -0.0198    0.1707  -0.12   0.91
## Year2009       -0.0749    0.1666  -0.45   0.65
## Year2010       -0.0113    0.1508  -0.07   0.94
## Year2011       -0.0631    0.1532  -0.41   0.68
## Year2012        0.1259    0.1531   0.82   0.41
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.536
## Multiple R-squared:  0.0346, Adjusted R-squared:  -0.00121
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 42 weights are ~= 1. The remaining 434 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.397  0.878  0.946  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      2.10e-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

```

```

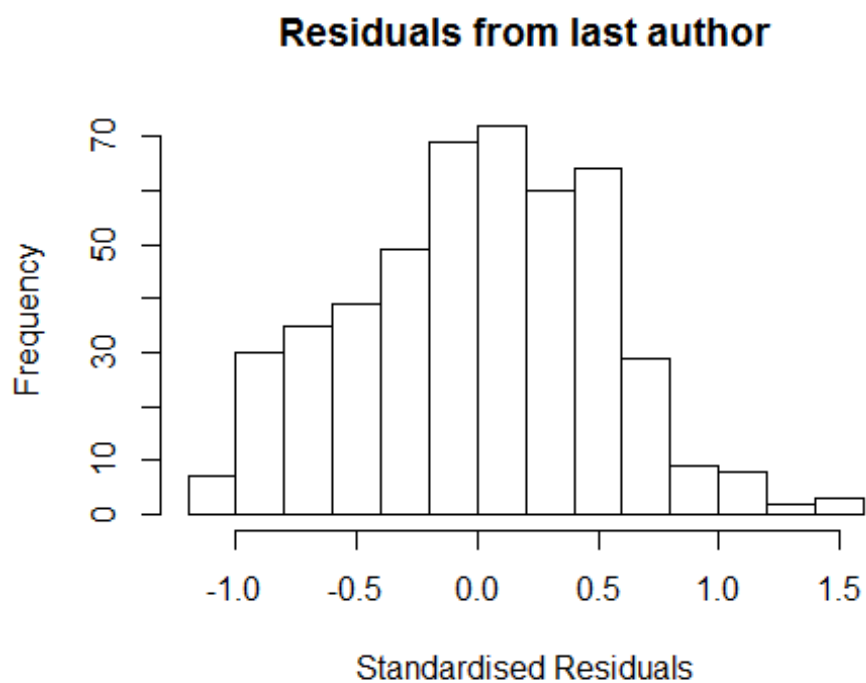
##                                "SM"
## seed : int(0)
## [1] "Sample size for the above analysis: 476"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3201"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    8    7    4    6    3    4    7    3    2    7    2    5    9    4
## 2011 2012
##    2    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    8    2    2    3    3    2    6    3    2    6    2    3    8    3
## 2011 2012
##    2    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    7    2    2    3    3    2    6    3    1    6    2    2    6    3
## 2011 2012
##    2    7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male first 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 = 3.5, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4"
## [1] "Male last author team size 2018 geometric mean: 2.6207413942089"

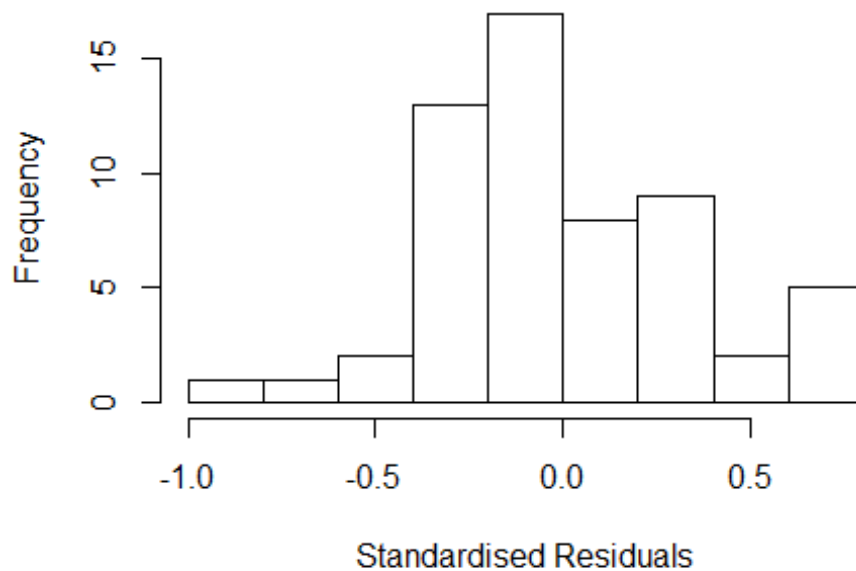
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, 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 2.994e+14  1      1.730e+07
## LastAuthorFemale  2.990e+13  1      5.468e+06
## UniqueAuthors    2.589e+32  4      1.126e+04
## Year              7.764e+33 16      1.146e+01
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -0.9394 -0.2367 -0.0258 0.2105 0.7590
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4051 0.3537 3.97 0.00034 ***
## FirstAuthorFemale1 0.3896 0.1673 2.33 0.02577 *
## LastAuthorFemale1 -0.1083 0.1394 -0.78 0.44252
## UniqueAuthors2 0.1763 0.2115 0.83 0.41013
## UniqueAuthors3 0.3187 0.1480 2.15 0.03823 *
## UniqueAuthors4 -0.3348 0.3313 -1.01 0.31913
## UniqueAuthors5 -0.0864 0.1982 -0.44 0.66554
## Year1997 -0.5678 0.3832 -1.48 0.14741
## Year1998 -0.7701 0.3872 -1.99 0.05455 .
## Year1999 -0.1128 0.3550 -0.32 0.75255
```

```

## Year2000          -0.8631      0.3824    -2.26  0.03035 *
## Year2001          -0.8722      0.4590    -1.90  0.06566 .
## Year2002          -0.1137      0.4093    -0.28  0.78280
## Year2003          -0.8585      0.3598    -2.39  0.02258 *
## Year2004          -0.7475      0.4909    -1.52  0.13683
## Year2005           0.3389      0.3537     0.96  0.34450
## Year2006          -0.6221      0.3676    -1.69  0.09948 .
## Year2007          -1.3499      0.3745    -3.61  0.00096 ***
## Year2008          -0.8146      0.3652    -2.23  0.03222 *
## Year2009          -0.7233      0.3455    -2.09  0.04363 *
## Year2010          -0.8368      0.4148    -2.02  0.05141 .
## Year2011          -1.2457      0.5899    -2.11  0.04191 *
## Year2012          -0.0776      0.3729    -0.21  0.83634
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.414
## Multiple R-squared:  0.609, Adjusted R-squared:  0.364
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 49 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.586  0.924  0.964  0.929  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          1.72e-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 3.540e+03 1      5.950e+01
## LastAuthorFemale  6.086e+13 1      7.801e+06
## Year              2.086e+17 16      3.477e+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)

```

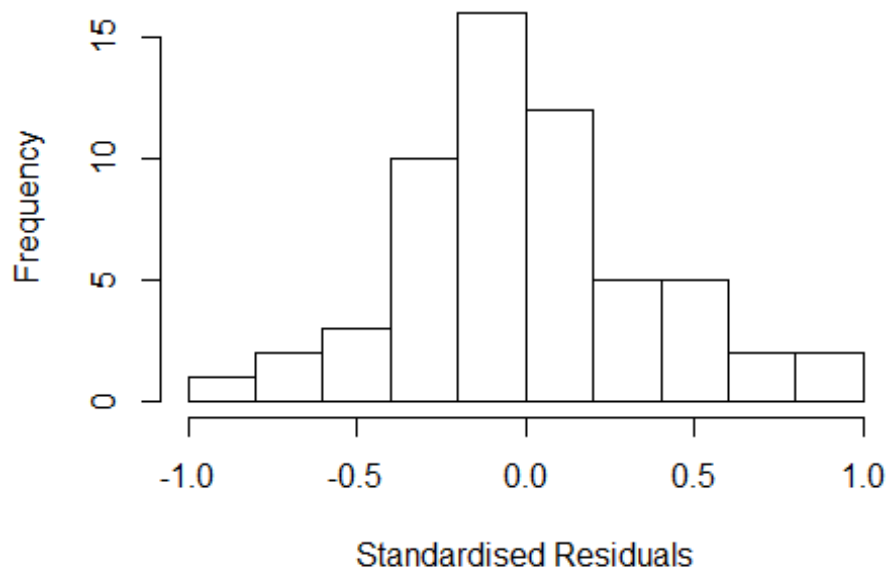
```
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##   control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.83895 -0.20928 -0.00456  0.17922  0.90643
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1286    0.1379   8.18 5.4e-10 ***
## FirstAuthorFemale1  0.4293    0.1565   2.74  0.0091 **
## LastAuthorFemale1 -0.1666    0.1379  -1.21  0.2344
## Year1997          -0.1509    0.1601  -0.94  0.3519
## Year1998          -0.4936    0.2093  -2.36  0.0235 *
## Year1999           0.4695    0.0544   8.63 1.4e-10 ***
## Year2000          -0.5797    0.1910  -3.03  0.0043 **
## Year2001          -0.5258    0.2781  -1.89  0.0661 .
## Year2002           0.2405    0.1819   1.32  0.1939
## Year2003          -0.4937    0.2343  -2.11  0.0416 *
## Year2004          -0.2406    0.2320  -1.04  0.3061
## Year2005           0.6154    0.1379   4.46 6.7e-05 ***
## Year2006          -0.2713    0.1801  -1.51  0.1399
## Year2007          -0.9759    0.1514  -6.45 1.2e-07 ***
## Year2008          -0.3694    0.3133  -1.18  0.2454
## Year2009          -0.5031    0.2465  -2.04  0.0481 *
## Year2010          -0.3978    0.2981  -1.33  0.1899
## Year2011          -0.7124    0.3811  -1.87  0.0691 .
## Year2012           0.2351    0.3326   0.71  0.4838
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.557, Adjusted R-squared:  0.353
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 48 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.613  0.890  0.960  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      1.72e-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"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

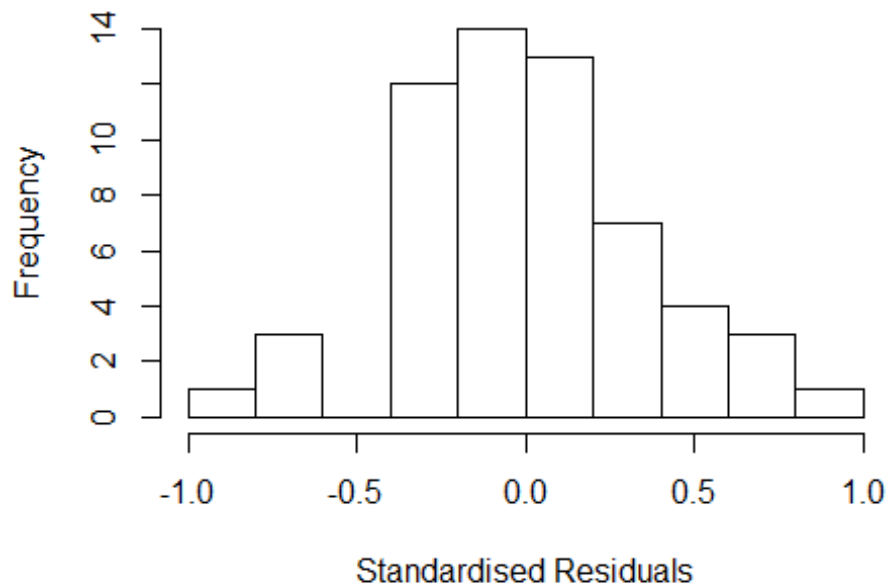
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1 NaN
## Year NaN 16 NaN
```

Residuals from first author



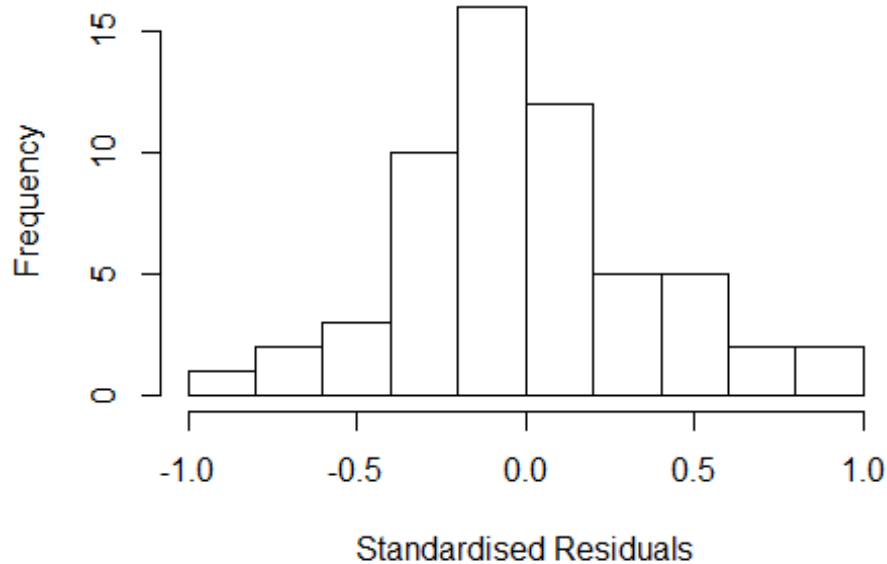
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.40e-01 -2.10e-01 -1.11e-15 2.04e-01 8.70e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9620 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 0.3423 0.1562 2.19 0.034 *
## Year1997 0.0170 0.1370 0.12 0.902
## Year1998 -0.3270 0.1567 -2.09 0.043 *
## Year1999 0.4695 0.0544 8.63 1.1e-10 ***
## Year2000 -0.4416 0.1782 -2.48 0.018 *
## Year2001 -0.3572 0.2303 -1.55 0.129
## Year2002 0.4107 0.1681 2.44 0.019 *
## Year2003 -0.3707 0.1652 -2.24 0.030 *
## Year2004 -0.1624 0.2906 -0.56 0.579
## Year2005 0.7820 0.0000 Inf < 2e-16 ***
## Year2006 -0.1344 0.1346 -1.00 0.324
```

```

## Year2007          -0.8491      0.1522    -5.58  1.9e-06 ***
## Year2008          -0.2426      0.3389    -0.72   0.478
## Year2009          -0.3343      0.1783    -1.88   0.068 .
## Year2010          -0.2276      0.3260    -0.70   0.489
## Year2011          -0.5856      0.3179    -1.84   0.073 .
## Year2012           0.3925      0.3385     1.16   0.253
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.438
## Multiple R-squared:  0.537, Adjusted R-squared:  0.34
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.673  0.927  0.973  0.930  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      1.72e-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 1.838e+14 1      1.356e+07
## Year              1.838e+14 16      2.791e+00

```

Residuals from last author



```
## [1] "List of 0 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
## -0.81679 -0.22975 -0.00156 0.20983 1.12822
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9336 0.1459 6.40 1.3e-07 ***
## LastAuthorFemale1 0.0284 0.1459 0.19 0.8465
## Year1997 0.2381 0.1722 1.38 0.1745
## Year1998 -0.2986 0.2140 -1.40 0.1706
## Year1999 0.4695 0.0544 8.63 1.1e-10 ***
## Year2000 -0.2988 0.1693 -1.76 0.0853 .
## Year2001 -0.3283 0.2782 -1.18 0.2449
## Year2002 0.7672 0.1027 7.47 4.2e-09 ***
## Year2003 -0.2862 0.2485 -1.15 0.2562
## Year2004 -0.0108 0.2614 -0.04 0.9673
## Year2005 0.8104 0.1459 5.55 2.0e-06 ***
## Year2006 -0.1109 0.1749 -0.63 0.5298
```

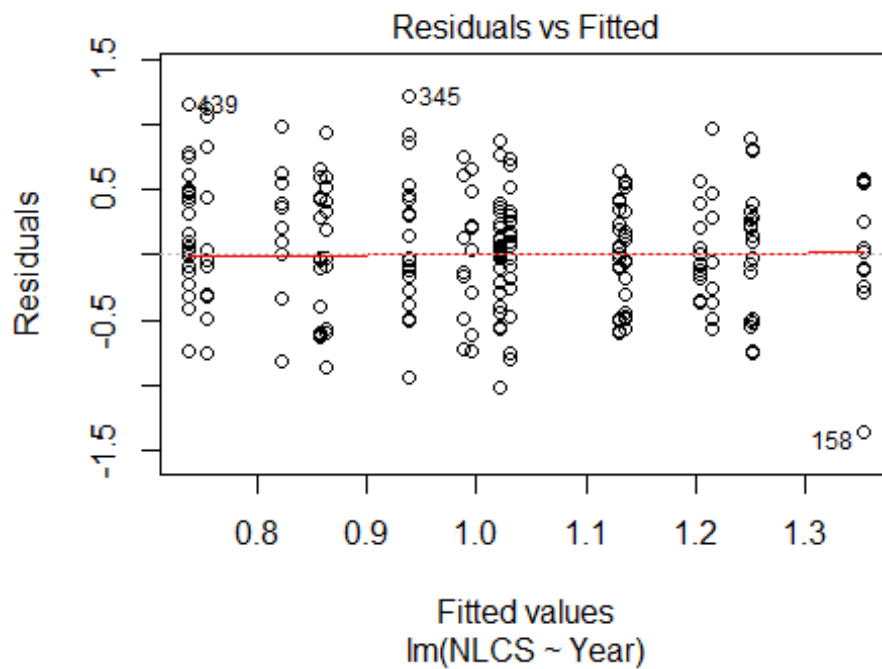


```

## Year2007          -0.6638      0.0746    -8.89  5.1e-11 ***
## Year2008          -0.0573      0.1927    -0.30  0.7678
## Year2009          -0.1504      0.2755    -0.55  0.5882
## Year2010          -0.0313      0.3889    -0.08  0.9364
## Year2011          -0.4003      0.5280    -0.76  0.4528
## Year2012           0.7118      0.2370     3.00  0.0046 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.495, Adjusted R-squared:  0.281
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 51 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.497  0.887  0.974  0.923  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      1.72e-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: 58"
## [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
##   15   11   11   15   12   17   12   12   13   20   19   20   18   19   29
## 2011 2012
##   36   39
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   14    8    7   12   11   11    8   12   11   18   16   18   13   16   24
## 2011 2012

```

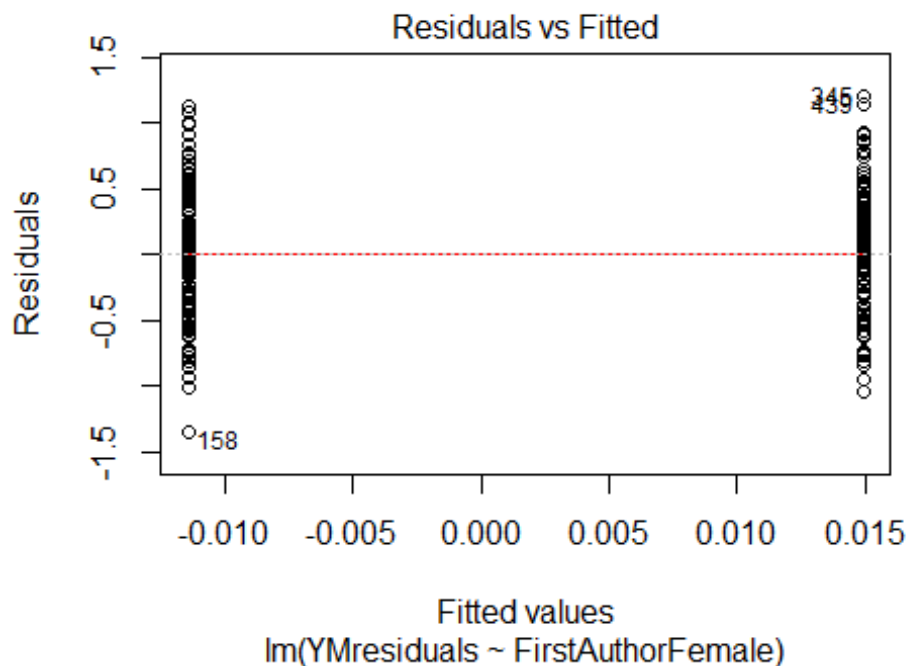
```
## 29 30
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 8 7 8 10 9 7 7 10 16 15 16 13 16 21
## 2011 2012
## 28 25
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.6
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.012, df = 1, p-value = 0.9
## [1] "Female first author team size 2018 geometric mean: 2.22074170124317"
## [1] "Male first author team size 2018 geometric mean: 2.08707383067606"
## 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] "Female last author team size 2018 geometric mean: 2.0718305416747"
## [1] "Male last author team size 2018 geometric mean: 2.30417316808542"

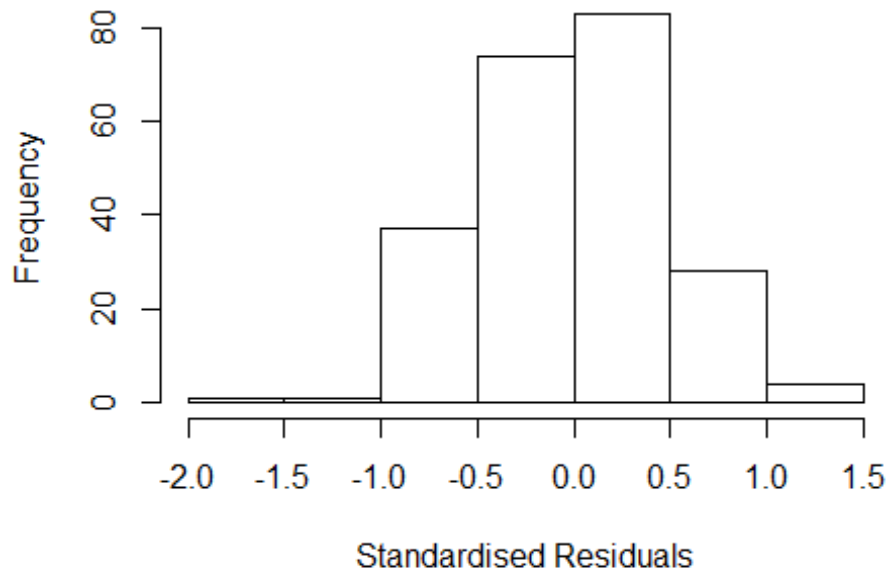
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 62, 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	1.661	1	1.289
LastAuthorFemale	1.564	1	1.251
UniqueAuthors	3.161	4	1.155
Year	5.851	16	1.057

Residuals from first and last author and team size



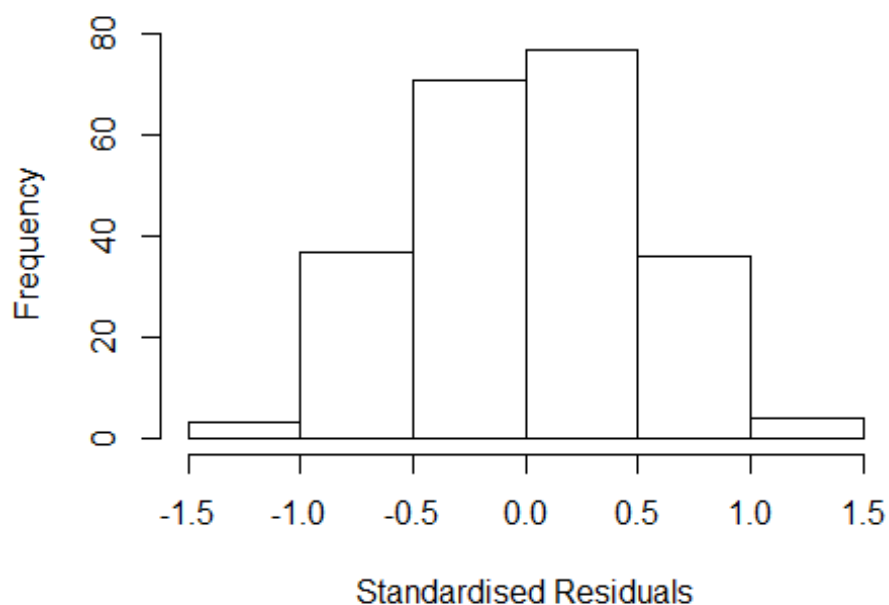
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.51094 -0.30770 0.00901 0.30587 1.26293
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7604 0.2399 3.17 0.0018 **
## FirstAuthorFemale1 -0.0236 0.0766 -0.31 0.7583
## LastAuthorFemale1 0.0327 0.0770 0.42 0.6717
## UniqueAuthors2 0.2001 0.0829 2.42 0.0166 *
## UniqueAuthors3 0.1572 0.1277 1.23 0.2198
## UniqueAuthors4 0.5162 0.1252 4.12 5.4e-05 ***
## UniqueAuthors5 0.3354 0.1433 2.34 0.0202 *
## Year1997 0.1437 0.2853 0.50 0.6151
## Year1998 0.1035 0.3951 0.26 0.7937
## Year1999 -0.1355 0.2873 -0.47 0.6377
```

```

## Year2000      0.3172      0.2604      1.22      0.2246
## Year2001      0.4754      0.2800      1.70      0.0911 .
## Year2002      0.1315      0.3047      0.43      0.6666
## Year2003      0.5177      0.2957      1.75      0.0814 .
## Year2004      0.1759      0.2629      0.67      0.5042
## Year2005      0.1006      0.2567      0.39      0.6955
## Year2006      0.2392      0.2512      0.95      0.3420
## Year2007      0.0987      0.2606      0.38      0.7054
## Year2008     -0.1544      0.3230     -0.48      0.6332
## Year2009     -0.0239      0.2633     -0.09      0.9279
## Year2010     -0.0309      0.2701     -0.11      0.9090
## Year2011      0.0715      0.2418      0.30      0.7677
## Year2012     -0.2144      0.2531     -0.85      0.3979
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.491
## Multiple R-squared:  0.189, Adjusted R-squared:  0.102
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 203 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.324  0.868  0.952  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      4.39e-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 1.642 1      1.282
## LastAuthorFemale  1.485 1      1.219
## Year              1.715 16      1.017

```

Residuals from first and last author



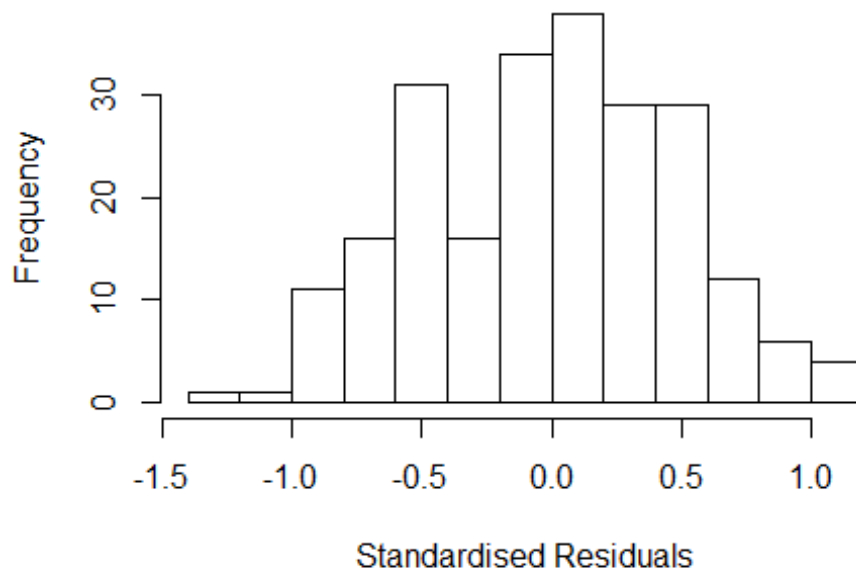
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.36185 -0.40220 0.00737 0.35692 1.19991
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.823723 0.232213 3.55 0.00048 ***
## FirstAuthorFemale1 -0.005641 0.082257 -0.07 0.94539
## LastAuthorFemale1 0.061289 0.079407 0.77 0.44108
## Year1997 0.173581 0.294556 0.59 0.55630
## Year1998 0.139173 0.318016 0.44 0.66211
## Year1999 -0.071066 0.312075 -0.23 0.82009
## Year2000 0.329613 0.252550 1.31 0.19328
## Year2001 0.548702 0.281647 1.95 0.05273 .
## Year2002 0.282178 0.310514 0.91 0.36453
## Year2003 0.476837 0.293852 1.62 0.10616
## Year2004 0.335909 0.245654 1.37 0.17297
## Year2005 0.277372 0.252969 1.10 0.27414
```

```

## Year2006          0.299156    0.247151    1.21  0.22748
## Year2007          0.139810    0.252340    0.55  0.58014
## Year2008         -0.154635    0.305877   -0.51  0.61371
## Year2009          0.009358    0.270939    0.03  0.97248
## Year2010          0.000195    0.260537    0.00  0.99940
## Year2011          0.149046    0.236195    0.63  0.52871
## Year2012         -0.108796    0.247300   -0.44  0.66044
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.116, Adjusted R-squared:  0.0399
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 198 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.489  0.878  0.940  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      4.39e-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 1.368 1      1.17
## Year              1.368 16      1.01

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3098 -0.4137  0.0205  0.3582  1.1870
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8402    0.2304   3.65 0.00033 ***
## FirstAuthorFemale1 0.0167    0.0755   0.22 0.82549
## Year1997        0.1651    0.2934   0.56 0.57428
## Year1998        0.1373    0.3160   0.43 0.66447
## Year1999       -0.0812    0.3148  -0.26 0.79677
## Year2000        0.3183    0.2507   1.27 0.20558
## Year2001        0.5540    0.2858   1.94 0.05394 .
## Year2002        0.2775    0.3086   0.90 0.36956
## Year2003        0.4697    0.2850   1.65 0.10090
## Year2004        0.3330    0.2452   1.36 0.17580
## Year2005        0.2704    0.2535   1.07 0.28722
## Year2006        0.2923    0.2474   1.18 0.23867
```

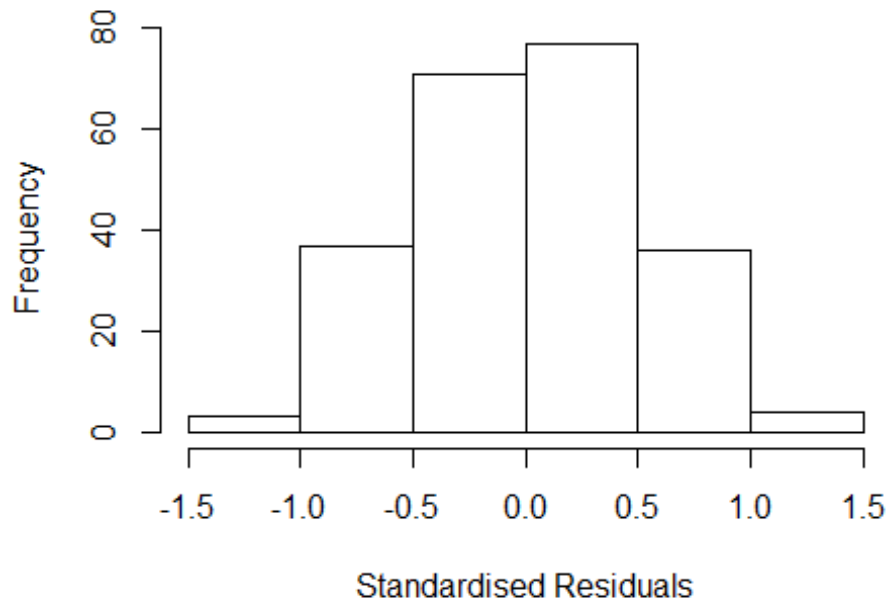


```

## Year2007          0.1445      0.2538      0.57  0.56986
## Year2008         -0.1581      0.3115     -0.51  0.61218
## Year2009          0.0200      0.2748      0.07  0.94208
## Year2010          0.0150      0.2623      0.06  0.95460
## Year2011          0.1574      0.2382      0.66  0.50952
## Year2012         -0.1024      0.2501     -0.41  0.68260
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.114, Adjusted R-squared:  0.0419
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 197 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.507  0.878  0.939  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      4.39e-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 1.232 1      1.110
## Year              1.232 16      1.007

```

Residuals from last author



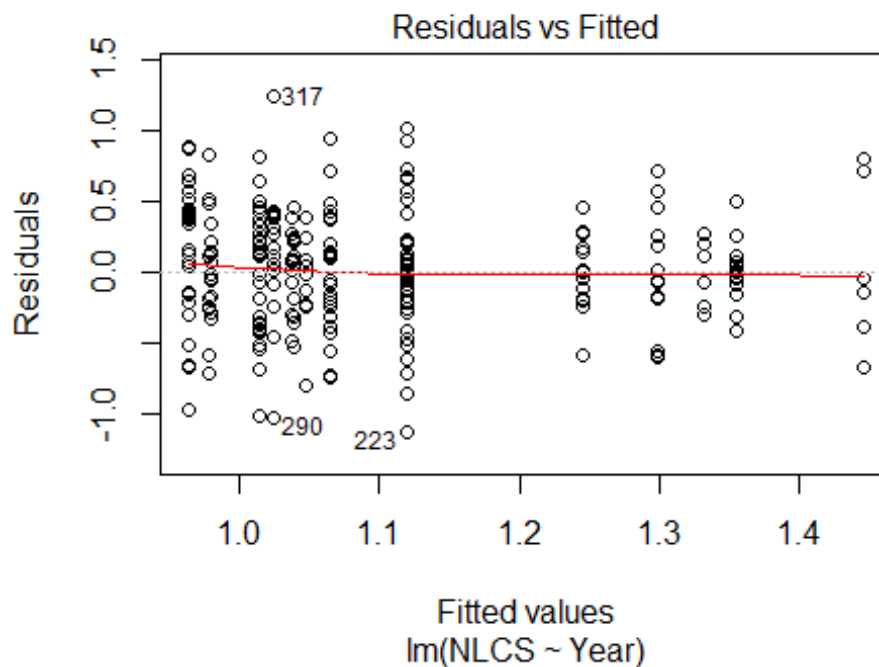
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.36449 -0.39585  0.00992  0.35821  1.20763
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.82358    0.23640   3.48  0.0006 ***
## LastAuthorFemale1 0.06060    0.07260   0.83  0.4049
## Year1997        0.17399    0.29689   0.59  0.5585
## Year1998        0.13823    0.32204   0.43  0.6682
## Year1999       -0.07428    0.31923  -0.23  0.8162
## Year2000        0.32854    0.25727   1.28  0.2030
## Year2001        0.54653    0.28570   1.91  0.0571 .
## Year2002        0.27687    0.31761   0.87  0.3844
## Year2003        0.48031    0.28892   1.66  0.0979 .
## Year2004        0.33468    0.25109   1.33  0.1840
## Year2005        0.27557    0.25849   1.07  0.2876
## Year2006        0.29660    0.25157   1.18  0.2397
```

```

## Year2007      0.13774      0.25722      0.54      0.5929
## Year2008     -0.16221      0.31712     -0.51      0.6095
## Year2009      0.00794      0.27664      0.03      0.9771
## Year2010     -0.00216      0.26624     -0.01      0.9935
## Year2011      0.14796      0.24097      0.61      0.5399
## Year2012     -0.11085      0.25253     -0.44      0.6612
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.513
## Multiple R-squared:  0.118, Adjusted R-squared:  0.0465
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 28 weights are ~= 1. The remaining 200 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.460  0.872  0.936  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      4.39e-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: 228"
## [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
##   13   15   12   16   12   13   13   17    7   15   16   22   20   20   37
## 2011 2012
##   34   31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   13    8   10    8    9   12   13    6   12   13   20   17   17   32
## 2011 2012

```

```
## 28 27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 13 6 9 6 9 11 12 6 12 11 19 16 16 31
## 2011 2012
## 28 27
## [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
```



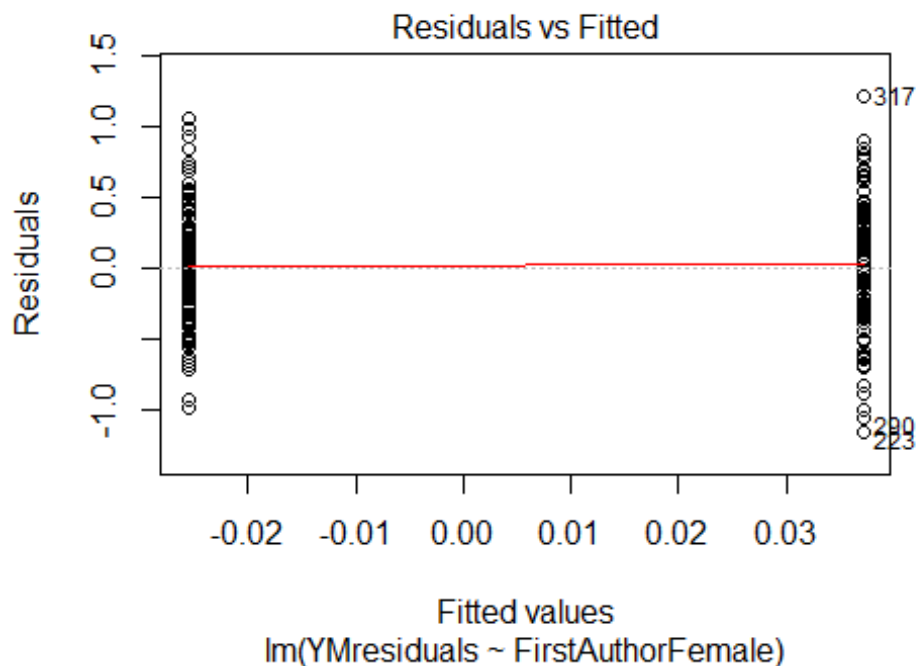
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.7, df = 1, p-value = 0.06

## [1] "Female first author team size 2018 geometric mean: 2.92773039467123"
## [1] "Male first author team size 2018 geometric mean: 3.32573393556655"

## 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 = 42, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.08618338496079"
## [1] "Male last author team size 2018 geometric mean: 2.99586201644542"

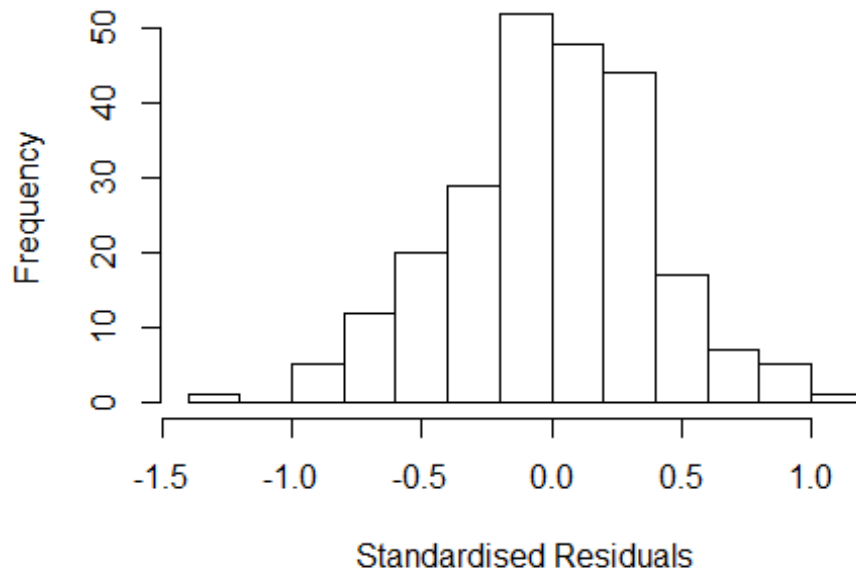
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 63, 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	1.406	1	1.186
LastAuthorFemale	1.269	1	1.127
UniqueAuthors	3.954	4	1.187
Year	5.268	16	1.053

Residuals from first and last author and team size



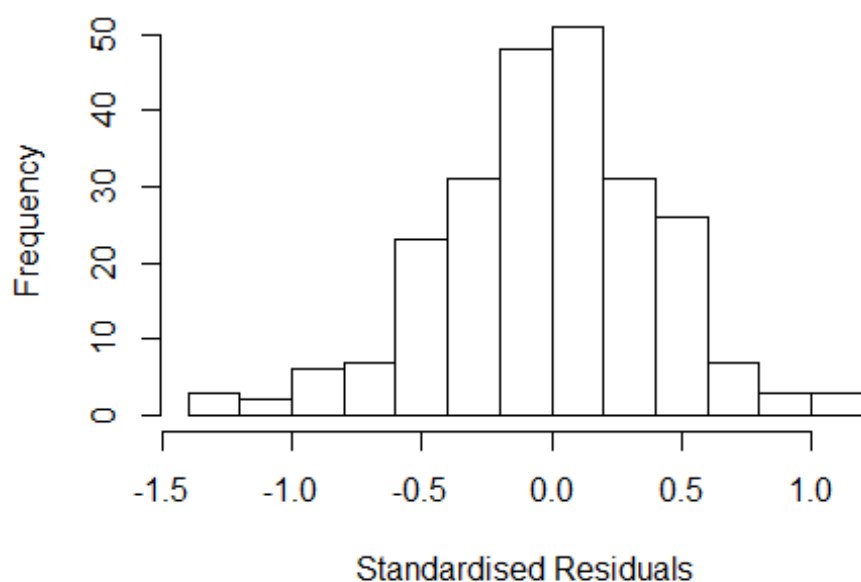
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.22881 -0.26771 0.00808 0.26186 1.01945
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8065 0.1148 7.02 2.7e-11 ***
## FirstAuthorFemale1 -0.1068 0.0580 -1.84 0.06712 .
## LastAuthorFemale1 -0.0884 0.0548 -1.61 0.10809
## UniqueAuthors2 0.2790 0.0849 3.29 0.00118 **
## UniqueAuthors3 0.2794 0.0895 3.12 0.00205 **
## UniqueAuthors4 0.4010 0.1111 3.61 0.00038 ***
## UniqueAuthors5 0.4392 0.0923 4.76 3.6e-06 ***
## Year1997 0.2726 0.1567 1.74 0.08334 .
## Year1998 0.1424 0.2274 0.63 0.53177
## Year1999 0.1124 0.1182 0.95 0.34252
```

```

## Year2000          0.3664      0.2085      1.76  0.08034 .
## Year2001          0.0236      0.1918      0.12  0.90211
## Year2002         -0.0052      0.1399     -0.04  0.97036
## Year2003          0.2546      0.1200      2.12  0.03502 *
## Year2004          0.3309      0.1258      2.63  0.00916 **
## Year2005          0.0148      0.1315      0.11  0.91052
## Year2006          0.4973      0.1345      3.70  0.00028 ***
## Year2007          0.1758      0.1486      1.18  0.23800
## Year2008          0.0981      0.1292      0.76  0.44840
## Year2009          0.1626      0.1557      1.04  0.29742
## Year2010          0.1281      0.1463      0.88  0.38247
## Year2011          0.1258      0.1344      0.94  0.35049
## Year2012          0.0881      0.1239      0.71  0.47785
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.392
## Multiple R-squared:  0.196, Adjusted R-squared:  0.115
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 217 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.305  0.876  0.953  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      4.15e-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 1.182 1      1.087
## LastAuthorFemale  1.275 1      1.129
## Year              1.500 16      1.013

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.24532 -0.26331 0.00117 0.24061 1.08937
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06326 0.10128 10.50 <2e-16 ***
## FirstAuthorFemale1 -0.05359 0.05606 -0.96 0.3401
## LastAuthorFemale1 -0.12690 0.05653 -2.24 0.0258 *
## Year1997 0.27898 0.15619 1.79 0.0754 .
## Year1998 0.13494 0.25891 0.52 0.6028
## Year1999 0.00802 0.13269 0.06 0.9519
## Year2000 0.31656 0.17332 1.83 0.0691 .
## Year2001 -0.04308 0.21475 -0.20 0.8412
## Year2002 0.05899 0.12663 0.47 0.6418
## Year2003 0.28509 0.12206 2.34 0.0204 *
## Year2004 0.37660 0.12921 2.91 0.0039 **
## Year2005 -0.01179 0.11299 -0.10 0.9170
```

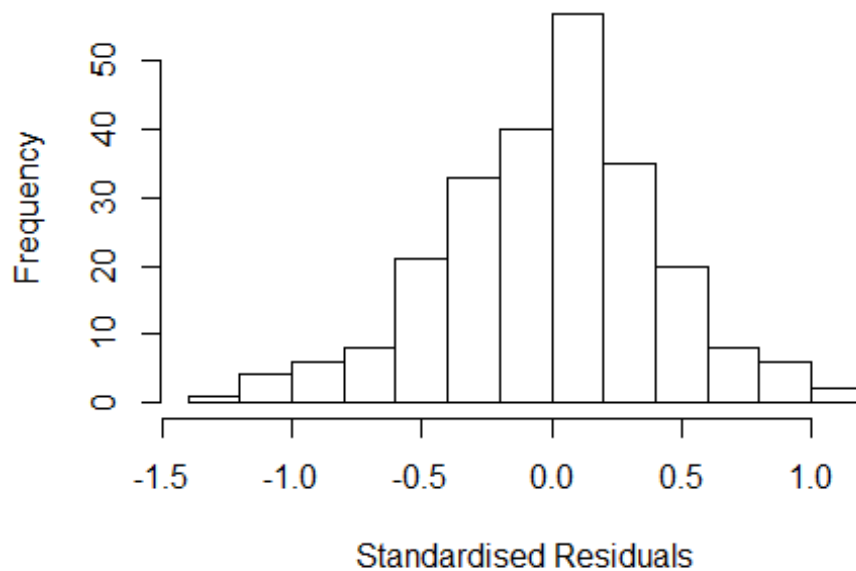


```

## Year2006          0.39037      0.11989      3.26      0.0013 **
## Year2007          0.18206      0.15071      1.21      0.2283
## Year2008          0.04196      0.12877      0.33      0.7448
## Year2009          0.16254      0.14966      1.09      0.2786
## Year2010          0.00713      0.16690      0.04      0.9659
## Year2011          0.06982      0.13322      0.52      0.6007
## Year2012          0.08261      0.13061      0.63      0.5277
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.397
## Multiple R-squared:  0.101, Adjusted R-squared:  0.0285
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 218 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.303  0.873   0.953   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      4.15e-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 1.232 1      1.110
## Year              1.232 16      1.007

```

Residuals from first author



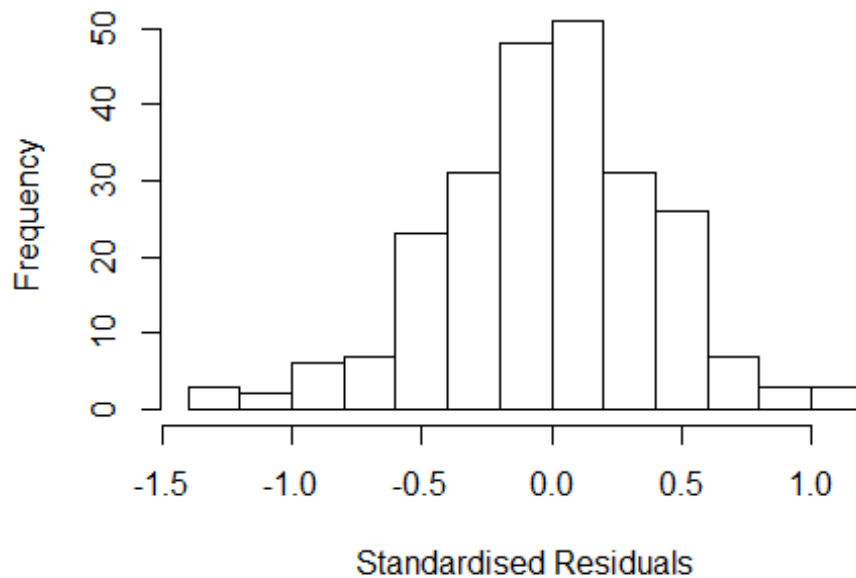
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2020 -0.2629 0.0274 0.2345 1.1340
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0439 0.1116 9.36 <2e-16 ***
## FirstAuthorFemale1 -0.0706 0.0584 -1.21 0.2283
## Year1997 0.2767 0.1646 1.68 0.0941 .
## Year1998 0.1187 0.2911 0.41 0.6839
## Year1999 0.0189 0.1442 0.13 0.8957
## Year2000 0.3041 0.1966 1.55 0.1233
## Year2001 -0.0430 0.2327 -0.18 0.8534
## Year2002 0.0220 0.1385 0.16 0.8739
## Year2003 0.2632 0.1303 2.02 0.0447 *
## Year2004 0.3478 0.1500 2.32 0.0213 *
## Year2005 -0.0344 0.1230 -0.28 0.7799
## Year2006 0.3517 0.1281 2.75 0.0065 **
```

```

## Year2007          0.1580      0.1500      1.05      0.2934
## Year2008          0.0336      0.1369      0.25      0.8062
## Year2009          0.1341      0.1600      0.84      0.4030
## Year2010         -0.0012      0.1770     -0.01      0.9946
## Year2011          0.0373      0.1433      0.26      0.7951
## Year2012          0.0641      0.1396      0.46      0.6467
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.396
## Multiple R-squared:  0.0806, Adjusted R-squared:  0.0105
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 216 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.336  0.852  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      4.15e-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 1.285 1      1.134
## Year            1.285 16      1.008

```

Residuals from last author



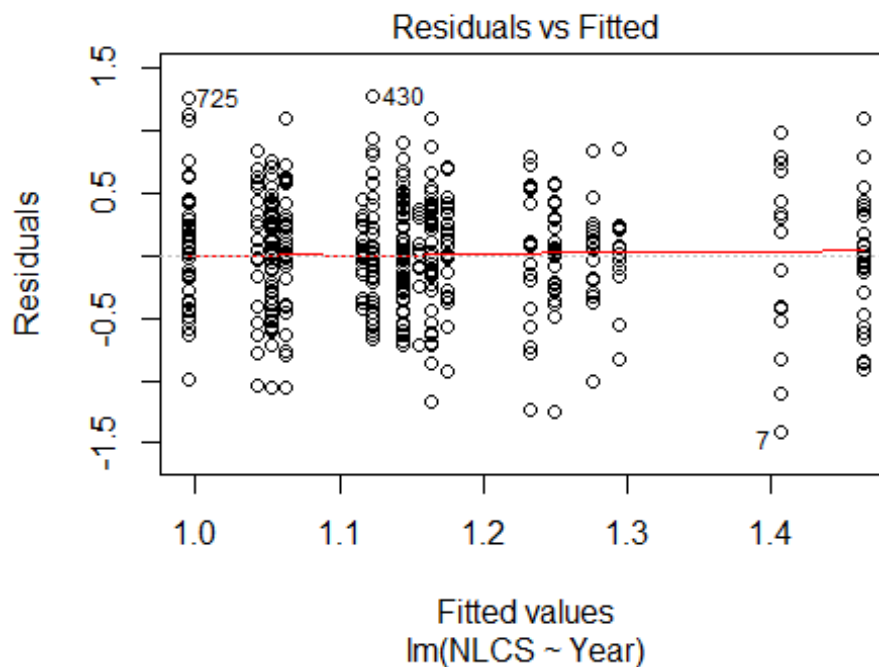
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.21019 -0.24823  0.00305  0.23717  1.07011
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.04040    0.09753   10.67  <2e-16 ***
## LastAuthorFemale1 -0.13433    0.05824   -2.31  0.0220 *
## Year1997         0.28307    0.16021    1.77  0.0786 .
## Year1998         0.13306    0.26206    0.51  0.6121
## Year1999         0.00775    0.12858    0.06  0.9520
## Year2000         0.29922    0.17497    1.71  0.0886 .
## Year2001        -0.05417    0.21468   -0.25  0.8010
## Year2002         0.05872    0.12735    0.46  0.6452
## Year2003         0.27976    0.12684    2.21  0.0284 *
## Year2004         0.35837    0.12517    2.86  0.0046 **
## Year2005        -0.00883    0.11344   -0.08  0.9380
## Year2006         0.37293    0.12056    3.09  0.0022 **
```

```

## Year2007      0.16979      0.14988      1.13      0.2585
## Year2008      0.03456      0.12740      0.27      0.7865
## Year2009      0.15749      0.15132      1.04      0.2991
## Year2010      0.00057      0.16917      0.00      0.9973
## Year2011      0.05875      0.13333      0.44      0.6599
## Year2012      0.07014      0.13127      0.53      0.5937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared:  0.0966, Adjusted R-squared:  0.0277
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 220 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.324  0.871  0.957  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      4.15e-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: 241"
## [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
##   18   27   21   26   31   27   28   15   25   22   32   28   42   43   56
## 2011 2012
##   74   59
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   16   26   19   18   22   12   24   12   23   16   27   27   35   35   48
## 2011 2012

```

```
## 65 48
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 25 13 17 19 11 19 11 22 14 25 24 28 30 44
## 2011 2012
## 59 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 = 36, df = 16, p-value = 0.003
```



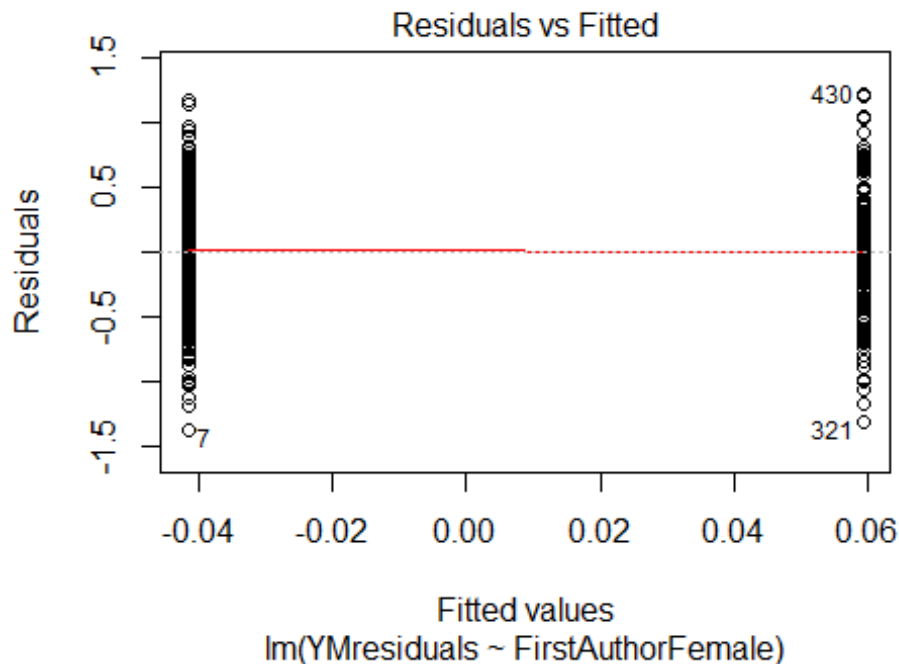
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.3, df = 1, p-value = 0.6

## [1] "Female first author team size 2018 geometric mean: 2.32403420726457"
## [1] "Male first author team size 2018 geometric mean: 2.4761105730299"

## 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 = 230, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.30049089467332"
## [1] "Male last author team size 2018 geometric mean: 2.50377411613655"

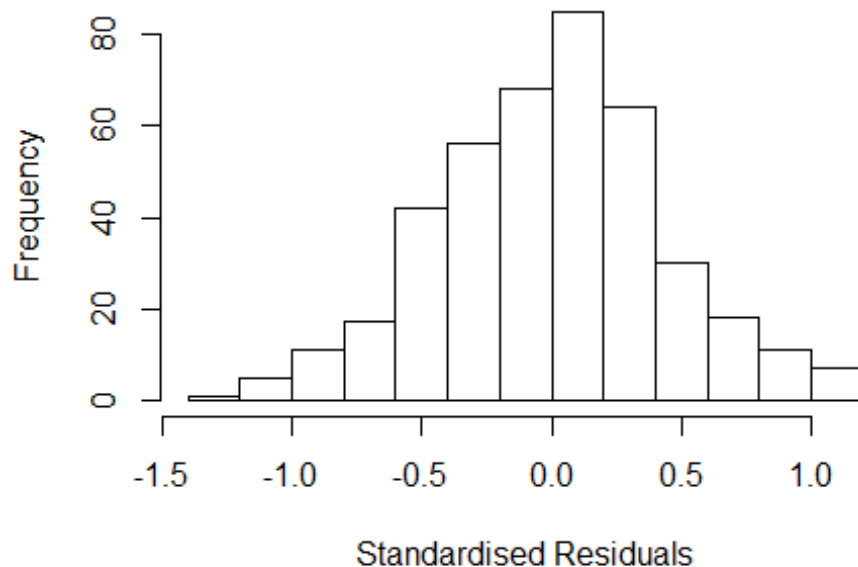
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 250, 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	1.407	1	1.186
LastAuthorFemale	1.332	1	1.154
UniqueAuthors	2.103	4	1.097
Year	2.532	16	1.029

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.2247 -0.2785  0.0211  0.2706  1.1808
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.10459    0.32240   3.43  0.00068 ***
## FirstAuthorFemale1 -0.04734    0.05055  -0.94  0.34958
## LastAuthorFemale1 -0.00532    0.04772  -0.11  0.91136
## UniqueAuthors2     0.17273    0.05740   3.01  0.00279 **
## UniqueAuthors3     0.22161    0.06884   3.22  0.00139 **
## UniqueAuthors4     0.33612    0.08992   3.74  0.00021 ***
## UniqueAuthors5     0.33356    0.10073   3.31  0.00101 **
## Year1997          0.23371    0.34463   0.68  0.49807
## Year1998         -0.16579    0.36708  -0.45  0.65178
## Year1999         -0.10948    0.32957  -0.33  0.73991
```

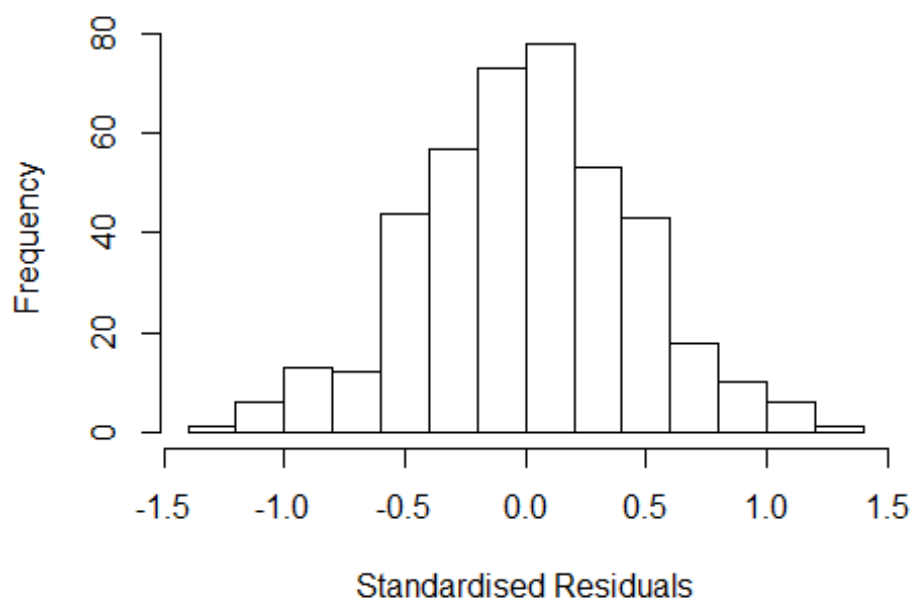


```

## Year2000      -0.11031      0.36426      -0.30      0.76217
## Year2001      0.00174      0.35300      0.00      0.99607
## Year2002     -0.09434      0.34787      -0.27      0.78640
## Year2003     -0.11148      0.33907      -0.33      0.74251
## Year2004     -0.07099      0.33889      -0.21      0.83419
## Year2005      0.03051      0.34754      0.09      0.93008
## Year2006      0.01050      0.33980      0.03      0.97537
## Year2007     -0.02239      0.33836      -0.07      0.94727
## Year2008     -0.18283      0.33661      -0.54      0.58733
## Year2009     -0.12135      0.33836      -0.36      0.72005
## Year2010     -0.07626      0.33567      -0.23      0.82039
## Year2011     -0.15873      0.32959      -0.48      0.63036
## Year2012     -0.27752      0.33392      -0.83      0.40643
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.415
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.0814
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 384 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.363  0.861  0.954  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-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 1.343 1      1.159
## LastAuthorFemale  1.278 1      1.131
## Year              1.327 16      1.009

```

Residuals from first and last author



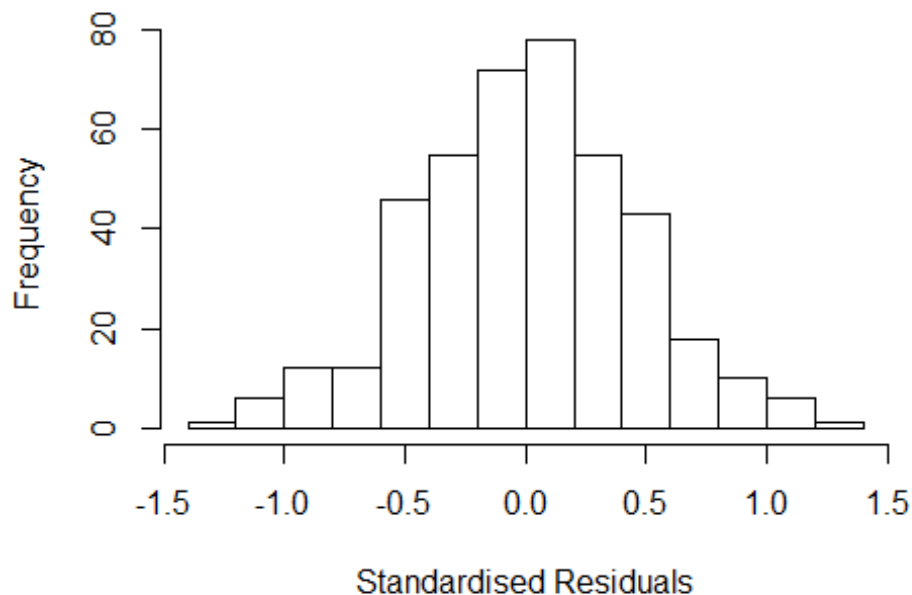
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.29602 -0.30603 0.00492 0.29309 1.21349
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.177514 0.355299 3.31 0.001 **
## FirstAuthorFemale1 -0.035087 0.050850 -0.69 0.491
## LastAuthorFemale1 -0.017308 0.047909 -0.36 0.718
## Year1997 0.299675 0.376289 0.80 0.426
## Year1998 -0.105859 0.385774 -0.27 0.784
## Year1999 -0.066475 0.359067 -0.19 0.853
## Year2000 -0.133343 0.394581 -0.34 0.736
## Year2001 0.122160 0.374894 0.33 0.745
## Year2002 -0.021100 0.371709 -0.06 0.955
## Year2003 0.013277 0.365441 0.04 0.971
## Year2004 -0.001368 0.367839 0.00 0.997
## Year2005 0.142579 0.371298 0.38 0.701
```

```

## Year2006          0.118507    0.362836    0.33    0.744
## Year2007          0.039915    0.367848    0.11    0.914
## Year2008         -0.121328    0.363201   -0.33    0.739
## Year2009         -0.029589    0.367099   -0.08    0.936
## Year2010          0.000325    0.363318    0.00    0.999
## Year2011         -0.127379    0.359171   -0.35    0.723
## Year2012         -0.190452    0.362331   -0.53    0.599
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.0731, Adjusted R-squared:  0.031
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 43 weights are ~= 1. The remaining 372 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.373  0.864  0.949  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      2.41e-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

```

Residuals from first author



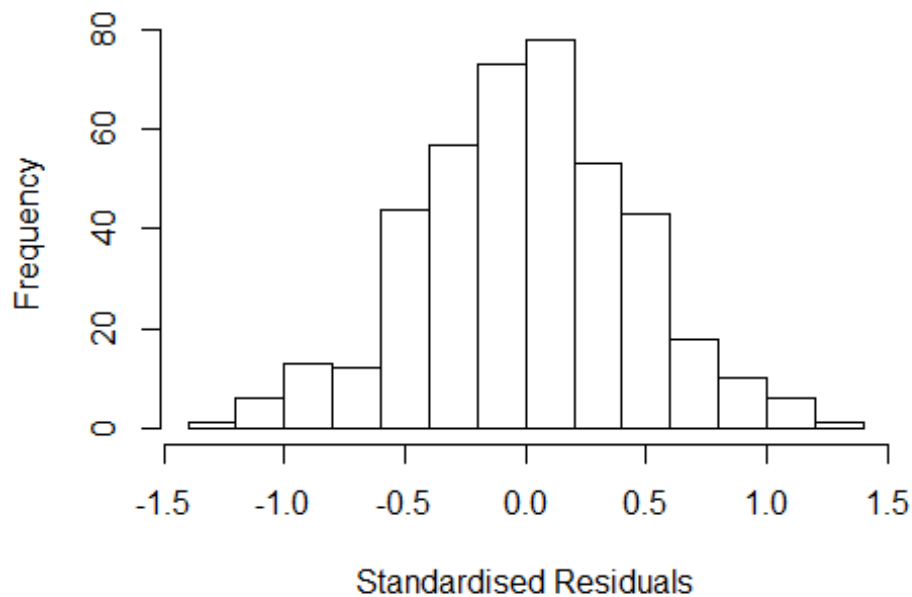
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2908 -0.3039  0.0107  0.2981  1.2164
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.174590   0.355274   3.31   0.001 **
## FirstAuthorFemale1 -0.042843   0.048001  -0.89   0.373
## Year1997         0.299520   0.377020   0.79   0.427
## Year1998        -0.108337   0.386802  -0.28   0.780
## Year1999        -0.066420   0.359649  -0.18   0.854
## Year2000        -0.136968   0.396030  -0.35   0.730
## Year2001         0.119195   0.375152   0.32   0.751
## Year2002        -0.022558   0.372633  -0.06   0.952
## Year2003         0.009670   0.366639   0.03   0.979
## Year2004        -0.000828   0.368417   0.00   0.998
## Year2005         0.139718   0.371863   0.38   0.707
## Year2006         0.116239   0.363747   0.32   0.749
```

```

## Year2007          0.037797    0.368754    0.10    0.918
## Year2008          -0.125001    0.364372   -0.34    0.732
## Year2009          -0.032691    0.368005   -0.09    0.929
## Year2010          -0.002499    0.364180   -0.01    0.995
## Year2011          -0.129133    0.359915   -0.36    0.720
## Year2012          -0.192037    0.363168   -0.53    0.597
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.444
## Multiple R-squared:  0.0727, Adjusted R-squared:  0.033
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 377 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.377  0.870  0.951  0.904  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      2.41e-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 1.126 1          1.061
## Year              1.126 16          1.004

```

Residuals from last author



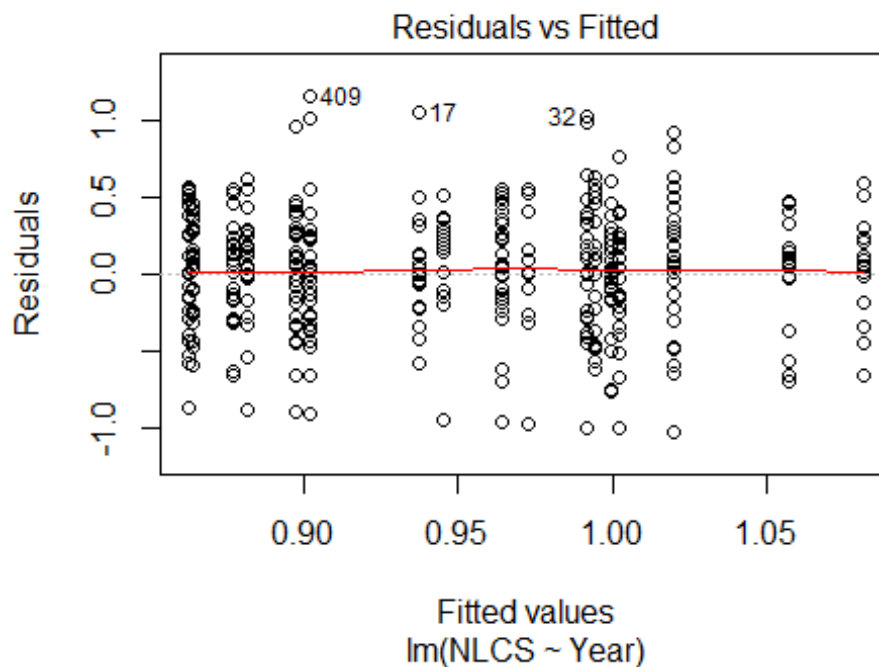
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28211 -0.30299  0.00877  0.28812  1.22047
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1705     0.3517   3.33 0.00096 ***
## LastAuthorFemale1 -0.0321     0.0453  -0.71 0.47799
## Year1997          0.2903     0.3715   0.78 0.43492
## Year1998         -0.1150     0.3810  -0.30 0.76297
## Year1999         -0.0686     0.3549  -0.19 0.84692
## Year2000         -0.1419     0.3904  -0.36 0.71649
## Year2001          0.1203     0.3711   0.32 0.74606
## Year2002         -0.0275     0.3665  -0.08 0.94022
## Year2003          0.0113     0.3612   0.03 0.97499
## Year2004         -0.0121     0.3627  -0.03 0.97332
## Year2005          0.1428     0.3677   0.39 0.69790
## Year2006          0.1116     0.3580   0.31 0.75545
```

```

## Year2007          0.0360      0.3638      0.10  0.92116
## Year2008          -0.1267      0.3589     -0.35  0.72423
## Year2009          -0.0323      0.3631     -0.09  0.92927
## Year2010          -0.0101      0.3583     -0.03  0.97758
## Year2011          -0.1347      0.3545     -0.38  0.70421
## Year2012          -0.1963      0.3580     -0.55  0.58376
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.0722, Adjusted R-squared:  0.0325
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 44 weights are ~= 1. The remaining 371 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.380  0.864  0.948  0.901  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      2.41e-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: 415"
## [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
##   24   28   24   24   24   24   27   29   27   18   34   31   26   32   37
## 2011 2012
##   38   36
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   22   24   20   19   16   17   25   24   25   12   33   30   22   25   27
## 2011 2012

```

```
## 28 25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 21 20 18 19 14 14 23 21 23 8 31 25 20 21 26
## 2011 2012
## 27 23
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.5
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.25, df = 1, p-value = 0.6

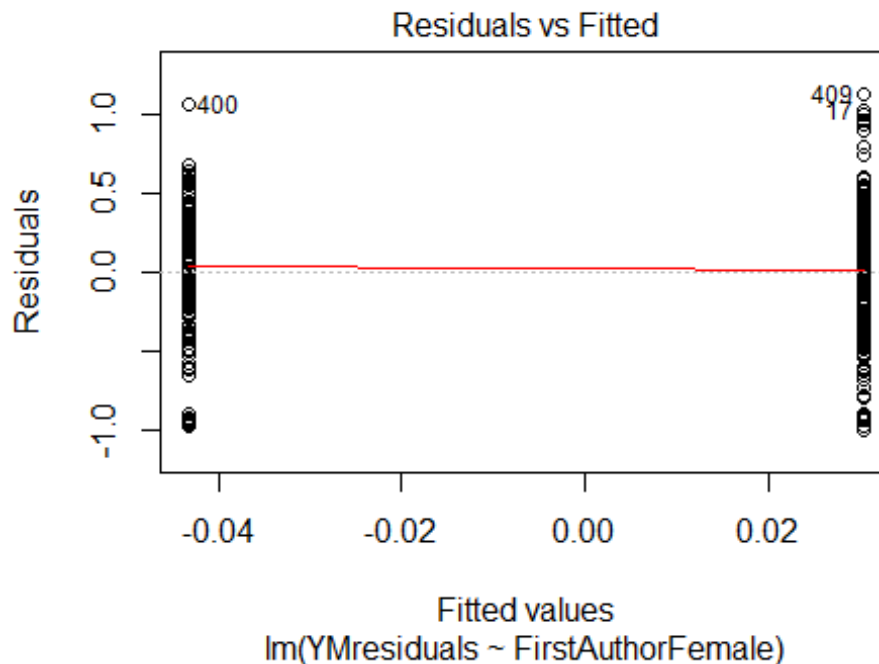
## [1] "Female first author team size 2018 geometric mean: 2.67734273546649"
## [1] "Male first author team size 2018 geometric mean: 1.97692818531885"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



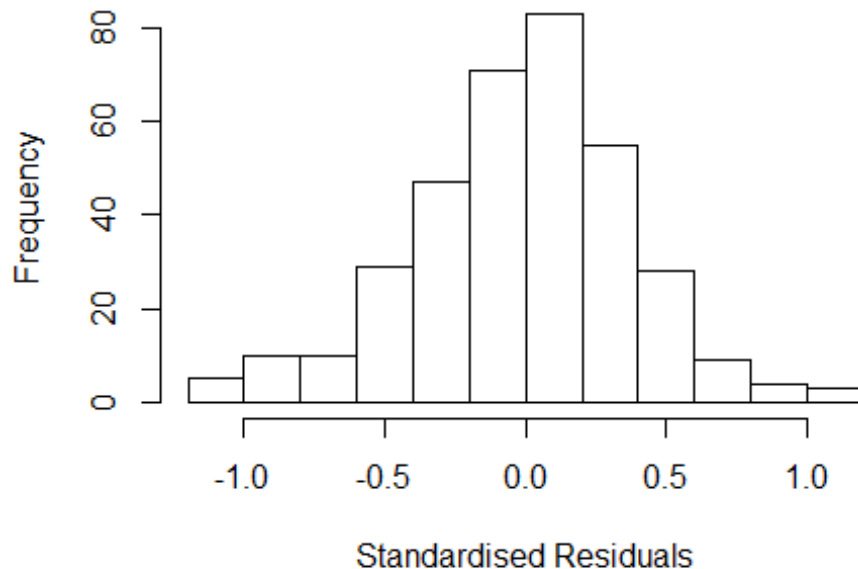
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 170, p-value = 0.2
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.70165143482398"
## [1] "Male last author team size 2018 geometric mean: 2.11401514822683"

## 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 = 0.4
## 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 1.498  1      1.224
## LastAuthorFemale  1.475  1      1.214
## UniqueAuthors    3.196  4      1.156
## Year              3.726 16      1.042
```

Residuals from first and last author and team size



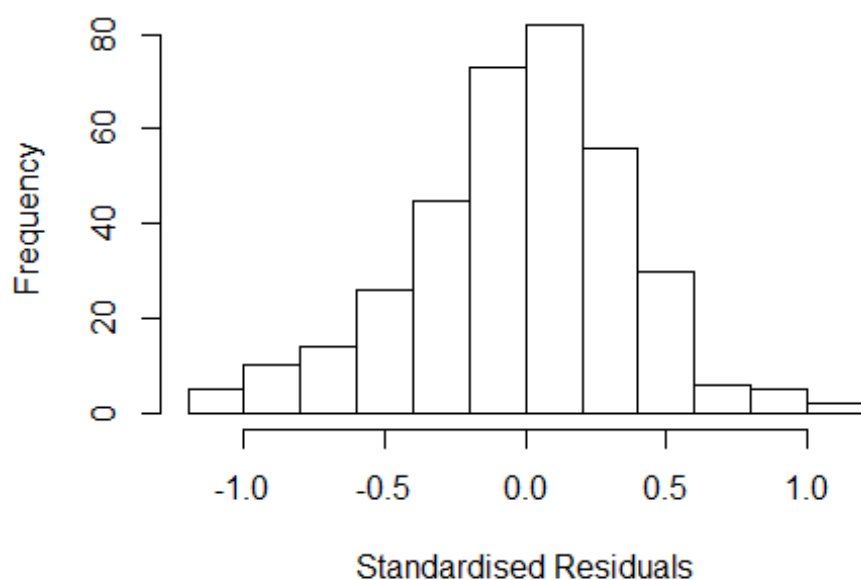
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0736 -0.2394 0.0139 0.2324 1.0784
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.90830 0.07983 11.38 <2e-16 ***
## FirstAuthorFemale1 -0.08419 0.04909 -1.71 0.087 .
## LastAuthorFemale1 0.06885 0.04991 1.38 0.169
## UniqueAuthors2 0.00837 0.06009 0.14 0.889
## UniqueAuthors3 0.07383 0.07346 1.01 0.316
## UniqueAuthors4 0.11547 0.07893 1.46 0.144
## UniqueAuthors5 0.21448 0.10239 2.09 0.037 *
## Year1997 0.03100 0.15697 0.20 0.844
## Year1998 0.09766 0.13594 0.72 0.473
## Year1999 0.12923 0.11401 1.13 0.258
```

```

## Year2000      0.08683    0.12714    0.68    0.495
## Year2001      0.19860    0.11739    1.69    0.092 .
## Year2002      0.09499    0.11071    0.86    0.392
## Year2003     -0.00582    0.10751   -0.05    0.957
## Year2004      0.14041    0.10267    1.37    0.172
## Year2005      0.16759    0.12912    1.30    0.195
## Year2006     -0.02302    0.10610   -0.22    0.828
## Year2007     -0.00856    0.10273   -0.08    0.934
## Year2008     -0.00685    0.09501   -0.07    0.943
## Year2009     -0.03088    0.09176   -0.34    0.737
## Year2010      0.13403    0.13489    0.99    0.321
## Year2011      0.05502    0.10896    0.50    0.614
## Year2012     -0.08286    0.11317   -0.73    0.465
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.368
## Multiple R-squared:  0.0567, Adjusted R-squared:  -0.00596
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 323 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.371  0.872  0.953  0.897  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      2.82e-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 1.355 1      1.164
## LastAuthorFemale  1.345 1      1.160
## Year              1.364 16      1.010

```

Residuals from first and last author



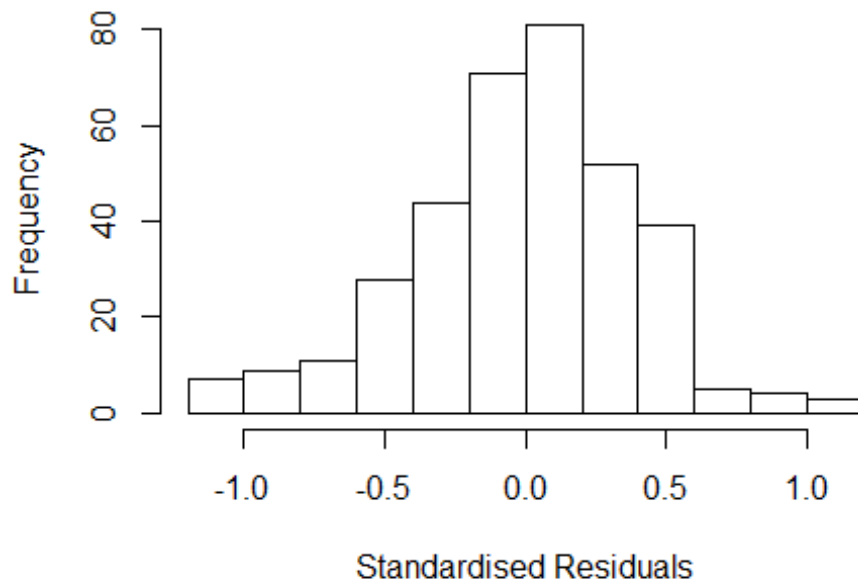
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0468 -0.2361 0.0105 0.2463 1.1029
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.922414 0.072375 12.74 <2e-16 ***
## FirstAuthorFemale1 -0.071717 0.047612 -1.51 0.133
## LastAuthorFemale1 0.060209 0.048620 1.24 0.216
## Year1997 0.085301 0.151577 0.56 0.574
## Year1998 0.088803 0.136506 0.65 0.516
## Year1999 0.124398 0.110704 1.12 0.262
## Year2000 0.109380 0.123527 0.89 0.377
## Year2001 0.229543 0.112592 2.04 0.042 *
## Year2002 0.109843 0.105071 1.05 0.297
## Year2003 0.039393 0.100432 0.39 0.695
## Year2004 0.138904 0.103556 1.34 0.181
## Year2005 0.188605 0.128833 1.46 0.144
```

```

## Year2006          0.000767    0.105429    0.01    0.994
## Year2007          0.026689    0.099429    0.27    0.789
## Year2008          0.025130    0.093242    0.27    0.788
## Year2009         -0.006781    0.095373   -0.07    0.943
## Year2010          0.157830    0.137683    1.15    0.252
## Year2011          0.105085    0.103989    1.01    0.313
## Year2012         -0.049810    0.114047   -0.44    0.663
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.363
## Multiple R-squared:  0.0427, Adjusted R-squared:  -0.00869
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 322 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.335  0.864  0.950  0.893  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      2.82e-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 1.17 1      1.082
## Year              1.17 16      1.005

```

Residuals from first author



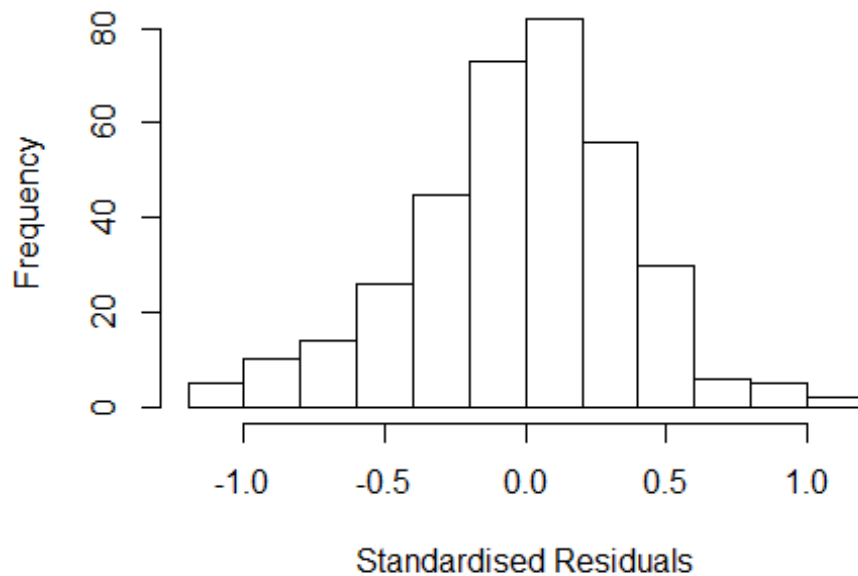
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0544 -0.2285 0.0166 0.2328 1.0951
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.927316 0.071062 13.05 <2e-16 ***
## FirstAuthorFemale1 -0.053274 0.044656 -1.19 0.234
## Year1997 0.108733 0.145835 0.75 0.456
## Year1998 0.089026 0.137826 0.65 0.519
## Year1999 0.127063 0.110575 1.15 0.251
## Year2000 0.130468 0.120265 1.08 0.279
## Year2001 0.233987 0.111065 2.11 0.036 *
## Year2002 0.119339 0.102391 1.17 0.245
## Year2003 0.047140 0.100274 0.47 0.639
## Year2004 0.139505 0.101493 1.37 0.170
## Year2005 0.182388 0.124975 1.46 0.145
## Year2006 0.014998 0.102420 0.15 0.884
```

```

## Year2007          0.029551    0.100312    0.29    0.768
## Year2008          0.026821    0.092432    0.29    0.772
## Year2009         -0.000146    0.093816    0.00    0.999
## Year2010          0.162313    0.134086    1.21    0.227
## Year2011          0.108893    0.103762    1.05    0.295
## Year2012         -0.040467    0.110201   -0.37    0.714
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.365
## Multiple R-squared:  0.0385, Adjusted R-squared:  -0.0101
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 36 weights are ~= 1. The remaining 318 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.348  0.865   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      2.82e-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 1.192 1      1.092
## Year              1.192 16      1.006

```

Residuals from last author



```
## [1] "List of 0 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
## -1.0695 -0.2490  0.0159  0.2278  1.1302
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.90295    0.07027   12.85  <2e-16 ***
## LastAuthorFemale1 0.03490    0.04530    0.77   0.442
## Year1997        0.08276    0.15043    0.55   0.583
## Year1998        0.08778    0.14233    0.62   0.538
## Year1999        0.13161    0.11103    1.19   0.237
## Year2000        0.12000    0.12039    1.00   0.320
## Year2001        0.23262    0.11015    2.11   0.035 *
## Year2002        0.11807    0.10179    1.16   0.247
## Year2003        0.05543    0.09932    0.56   0.577
## Year2004        0.13568    0.10293    1.32   0.188
## Year2005        0.17425    0.12605    1.38   0.168
## Year2006       -0.00165    0.10513   -0.02   0.988
```

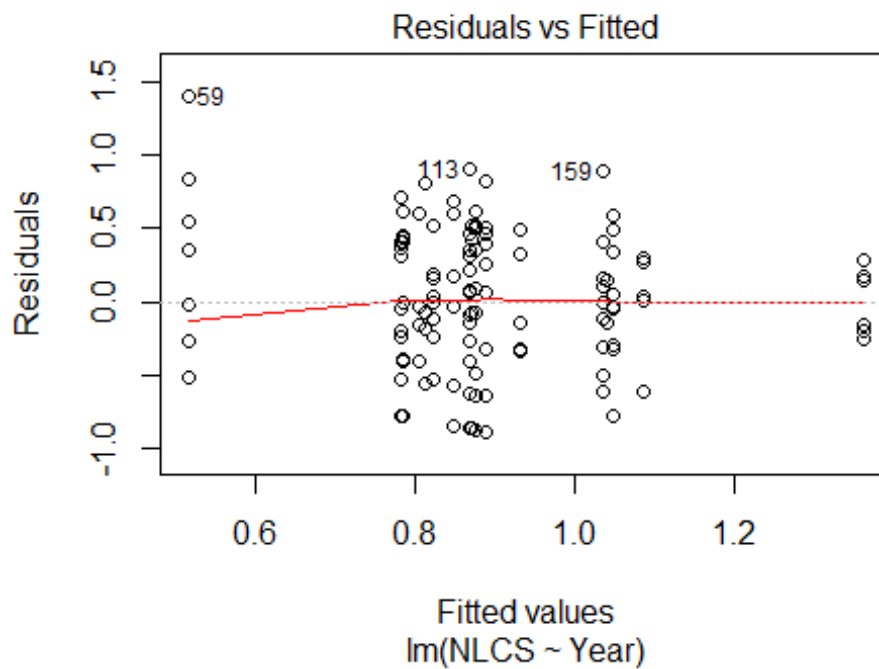


```

## Year2007          0.01883      0.10063      0.19      0.852
## Year2008          0.01646      0.09597      0.17      0.864
## Year2009         -0.01863      0.09591     -0.19      0.846
## Year2010          0.15436      0.13784      1.12      0.264
## Year2011          0.10130      0.10538      0.96      0.337
## Year2012         -0.05600      0.11601     -0.48      0.630
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.362
## Multiple R-squared:  0.0368, Adjusted R-squared:  -0.012
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 39 weights are ~= 1. The remaining 315 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.310  0.866  0.949  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      2.82e-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: 354"
## [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
##    2    5    6    6    5    6    9   12   11   10   12   13   10    5    9
## 2011 2012
##    5   10
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5    6    4    4    6    9   12   10    9   10   13    9    4    9
## 2011 2012

```

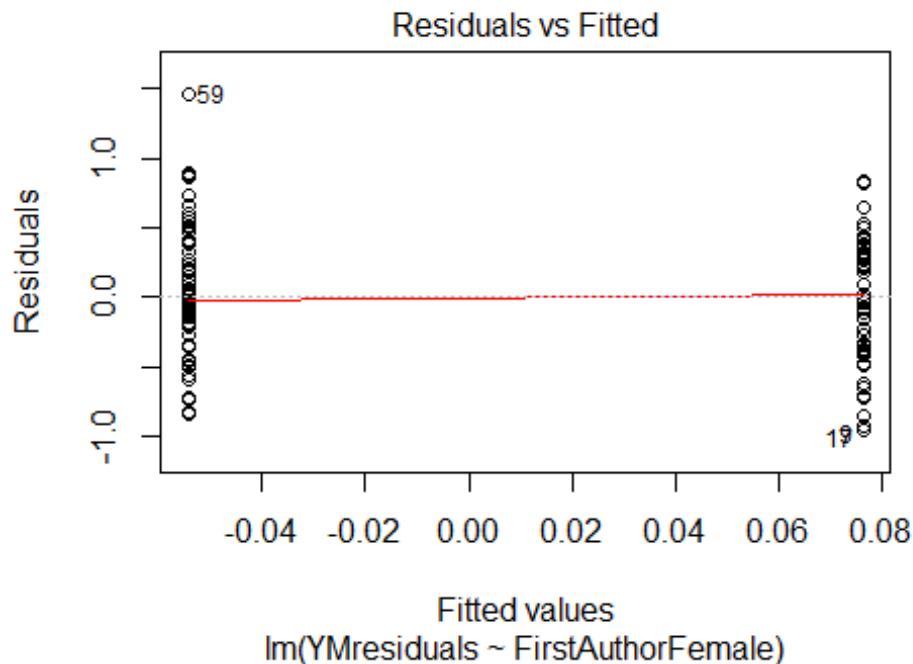
```
##      5      8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      3      6      4      4      6      8     12     10      7     10     10      9      4      8
## 2011 2012
##      5      7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 11, df = 16, p-value = 0.8
```



```
##
## 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"
## [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 = 2, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1"
## [1] "Male last author team size 2018 geometric mean: 2.51984209978975"

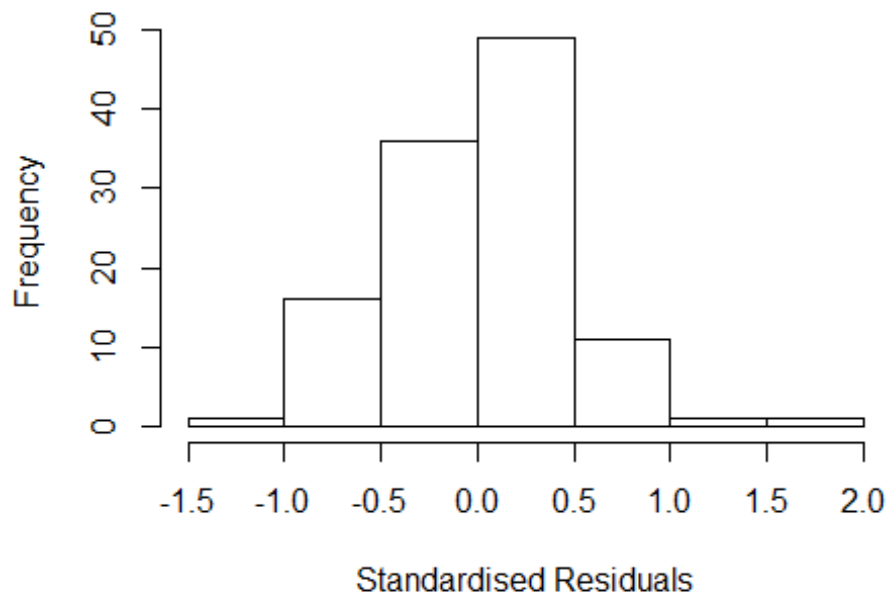
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = 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	3.043	1	1.744
LastAuthorFemale	1.852	1	1.361
UniqueAuthors	35.162	4	1.560
Year	47.870	16	1.129

Residuals from first and last author and team size



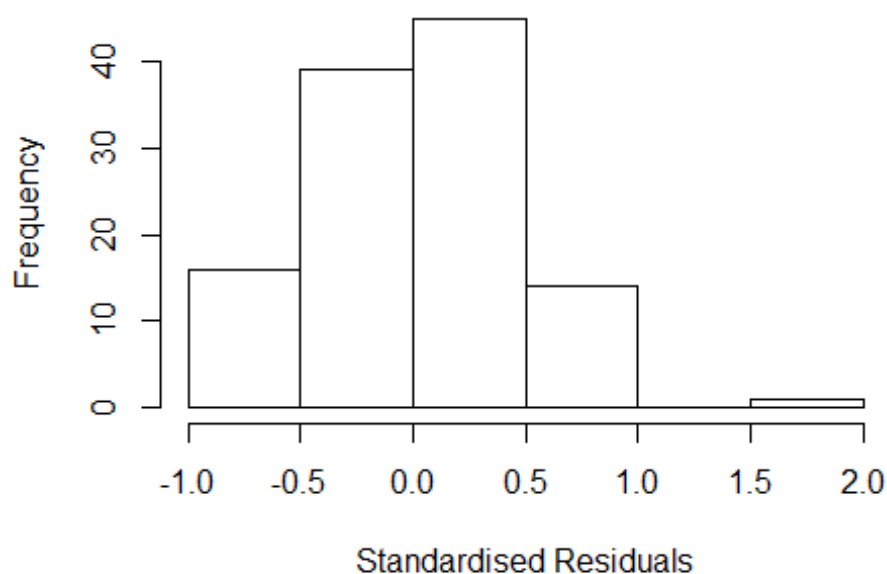
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4505 -0.2811 0.0163 0.3018 1.6379
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.04000 0.10160 10.24 < 2e-16 ***
## FirstAuthorFemale1 0.07378 0.12314 0.60 0.5505
## LastAuthorFemale1 -0.04278 0.11587 -0.37 0.7128
## UniqueAuthors2 0.20521 0.15525 1.32 0.1895
## UniqueAuthors3 0.54102 0.22879 2.36 0.0201 *
## UniqueAuthors4 0.38567 0.20804 1.85 0.0670 .
## UniqueAuthors5 0.33818 0.22749 1.49 0.1406
## Year1997 -0.21020 0.28027 -0.75 0.4552
## Year1998 -0.30425 0.36814 -0.83 0.4107
## Year1999 -0.20433 0.32179 -0.63 0.5270
```

```

## Year2000      -0.48274    0.28398   -1.70    0.0925 .
## Year2001      0.00479    0.17081    0.03    0.9777
## Year2002     -0.19995    0.24007   -0.83    0.4071
## Year2003     -0.75888    0.15669   -4.84    5.1e-06 ***
## Year2004     -0.42197    0.23331   -1.81    0.0738 .
## Year2005     -0.46835    0.21556   -2.17    0.0324 *
## Year2006     -0.29701    0.22349   -1.33    0.1871
## Year2007     -0.36373    0.22595   -1.61    0.1109
## Year2008     -0.34147    0.19868   -1.72    0.0890 .
## Year2009     -0.56888    0.20386   -2.79    0.0064 **
## Year2010     -0.30309    0.30099   -1.01    0.3166
## Year2011     -0.44760    0.18087   -2.47    0.0152 *
## Year2012     -0.40000    0.20560   -1.95    0.0548 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.442
## Multiple R-squared:  0.297, Adjusted R-squared:  0.129
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 103 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.140  0.869  0.952  0.903  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      8.70e-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 1.610 1      1.269
## LastAuthorFemale  1.757 1      1.325
## Year              2.267 16      1.026

```

Residuals from first and last author



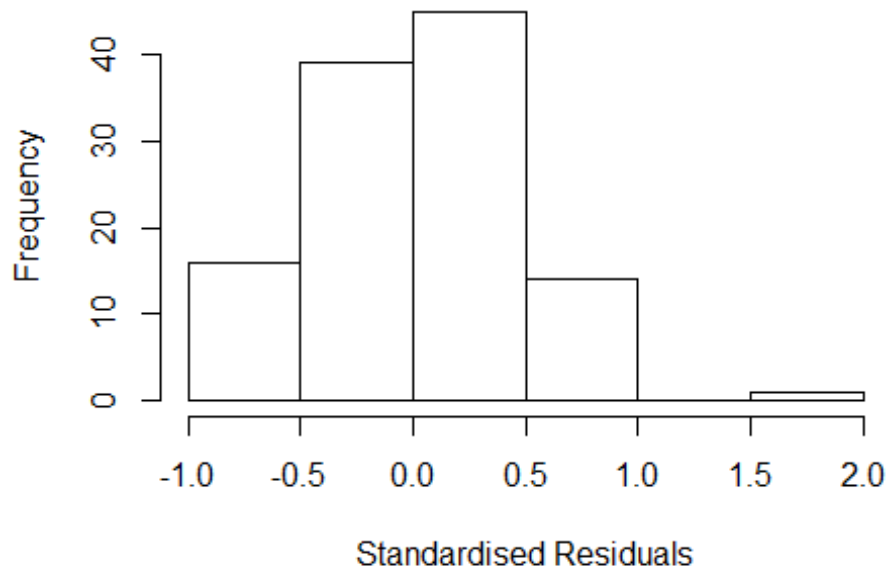
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9513 -0.3671 0.0171 0.3259 1.5287
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0400 0.1011 10.28 <2e-16 ***
## FirstAuthorFemale1 0.1877 0.1006 1.87 0.0650 .
## LastAuthorFemale1 -0.0297 0.1151 -0.26 0.7972
## Year1997 -0.1012 0.2291 -0.44 0.6597
## Year1998 -0.2467 0.3401 -0.73 0.4700
## Year1999 -0.2949 0.3408 -0.87 0.3890
## Year2000 -0.2369 0.3127 -0.76 0.4506
## Year2001 0.2276 0.1378 1.65 0.1019
## Year2002 -0.1211 0.2254 -0.54 0.5923
## Year2003 -0.6497 0.2088 -3.11 0.0025 **
## Year2004 -0.2727 0.2188 -1.25 0.2157
## Year2005 -0.3869 0.2223 -1.74 0.0850 .
```

```

## Year2006          -0.2346      0.2294   -1.02   0.3090
## Year2007          -0.1064      0.1561   -0.68   0.4973
## Year2008          -0.0811      0.1792   -0.45   0.6517
## Year2009          -0.3763      0.2220   -1.70   0.0932 .
## Year2010          -0.1088      0.1998   -0.54   0.5873
## Year2011          -0.1833      0.2331   -0.79   0.4334
## Year2012          -0.2099      0.1633   -1.29   0.2018
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.504
## Multiple R-squared:  0.184, Adjusted R-squared:  0.0311
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 103 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.337 0.891 0.946 0.918 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      8.70e-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 1.416 1      1.190
## Year      1.416 16      1.011

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.9672 -0.3817 0.0225 0.3194 1.5307
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0400 0.1012 10.28 <2e-16 ***
## FirstAuthorFemale1 0.1787 0.0948 1.89 0.0624 .
## Year1997 -0.1081 0.2238 -0.48 0.6303
## Year1998 -0.2515 0.3491 -0.72 0.4731
## Year1999 -0.2946 0.3446 -0.85 0.3948
## Year2000 -0.2678 0.2905 -0.92 0.3589
## Year2001 0.2321 0.1370 1.69 0.0934 .
## Year2002 -0.1251 0.2249 -0.56 0.5795
## Year2003 -0.6517 0.2096 -3.11 0.0025 **
## Year2004 -0.2773 0.2218 -1.25 0.2141
## Year2005 -0.3922 0.2221 -1.77 0.0806 .
## Year2006 -0.2326 0.2295 -1.01 0.3134
```

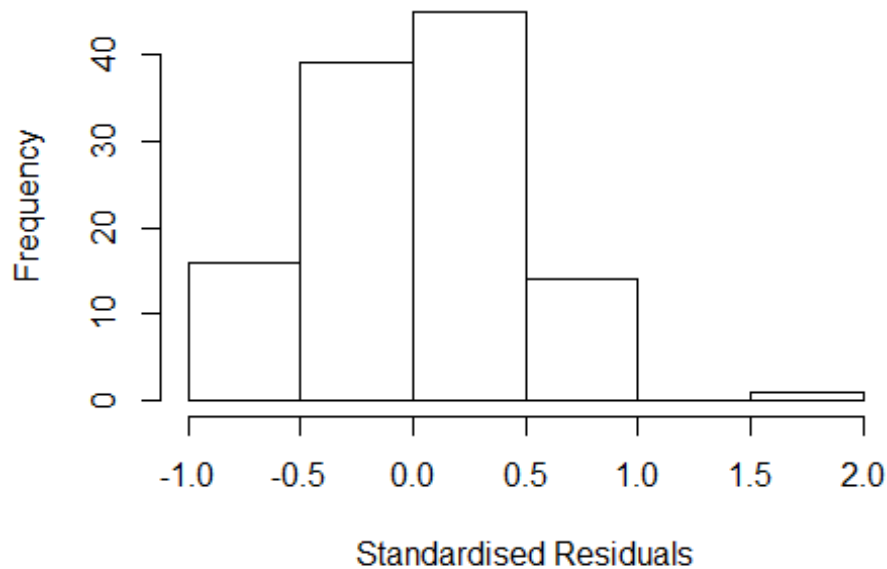


```

## Year2007          -0.1132      0.1516   -0.75    0.4571
## Year2008          -0.0845      0.1818   -0.46    0.6433
## Year2009          -0.3787      0.2205   -1.72    0.0891 .
## Year2010          -0.1175      0.1986   -0.59    0.5554
## Year2011          -0.1914      0.2313   -0.83    0.4099
## Year2012          -0.2229      0.1621   -1.38    0.1723
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.498
## Multiple R-squared:  0.185, Adjusted R-squared:  0.0424
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 104 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.324  0.886  0.945  0.917  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      8.70e-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 1.479 1          1.216
## Year            1.479 16          1.012

```

Residuals from last author



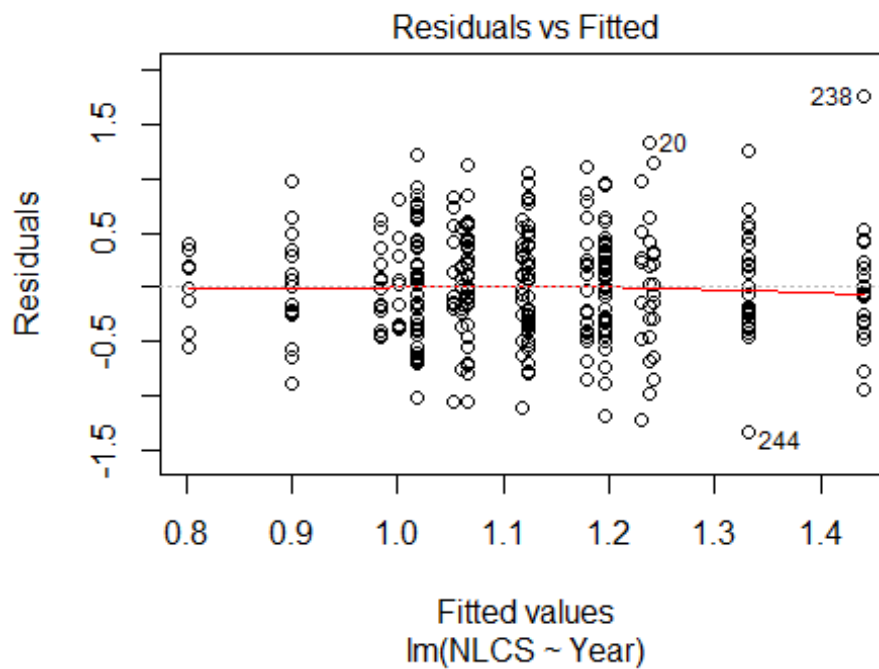
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9849 -0.3401 -0.0188  0.3821  1.5136
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.04000    0.10121   10.28  <2e-16 ***
## LastAuthorFemale1 0.04195    0.10800    0.39  0.6986
## Year1997       -0.06153    0.25148   -0.24  0.8072
## Year1998       -0.18556    0.30003   -0.62  0.5377
## Year1999       -0.12976    0.34805   -0.37  0.7101
## Year2000       -0.31044    0.31083   -1.00  0.3204
## Year2001        0.32003    0.13410    2.39  0.0189 *
## Year2002       -0.05510    0.22409   -0.25  0.8063
## Year2003       -0.63458    0.21826   -2.91  0.0045 **
## Year2004       -0.24189    0.23047   -1.05  0.2966
## Year2005       -0.28957    0.23511   -1.23  0.2211
## Year2006       -0.13644    0.23544   -0.58  0.5636
```

```

## Year2007          -0.03209      0.16141    -0.20    0.8428
## Year2008           0.00378      0.17689     0.02    0.9830
## Year2009          -0.26397      0.22870    -1.15    0.2512
## Year2010          -0.02607      0.20400    -0.13    0.8986
## Year2011          -0.13296      0.20060    -0.66    0.5090
## Year2012          -0.20023      0.15562    -1.29    0.2013
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.163, Adjusted R-squared:  0.0167
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 104 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.328  0.875   0.942   0.914   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      8.70e-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: 115"
## [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
##   11   13   16    9   12   18   12   19   13   19   24   27   32   40   49
## 2011 2012
##   52   56
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11   10   10    8    8   13   10   17   12   16   21   25   22   36   43
## 2011 2012

```

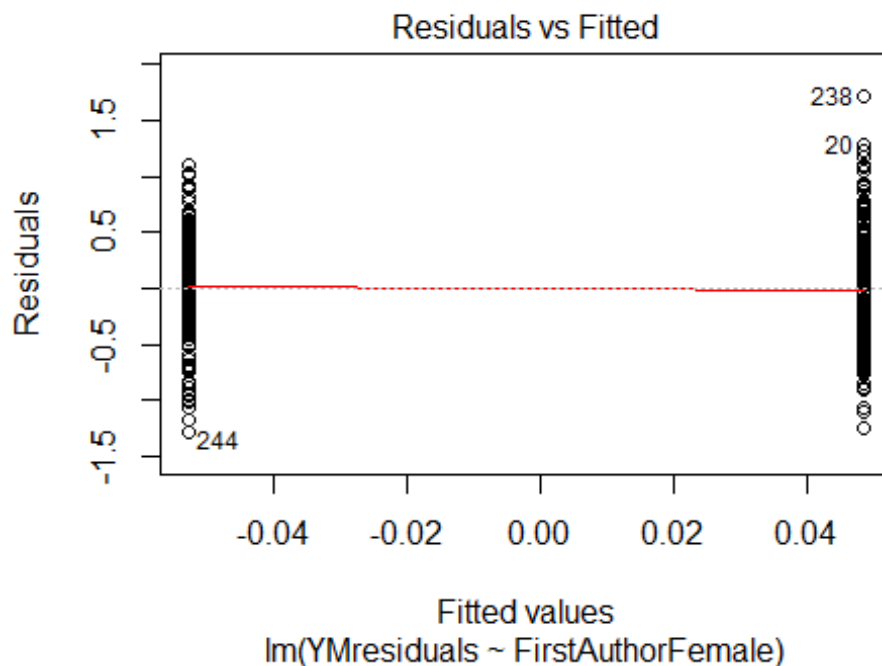
```
## 46 43
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 9 10 10 8 8 11 9 16 11 16 21 22 19 35 36
## 2011 2012
## 43 39
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.7
```



```
##
## 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: 2.02153015930144"
## [1] "Male first author team size 2018 geometric mean: 2.39754115229153"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

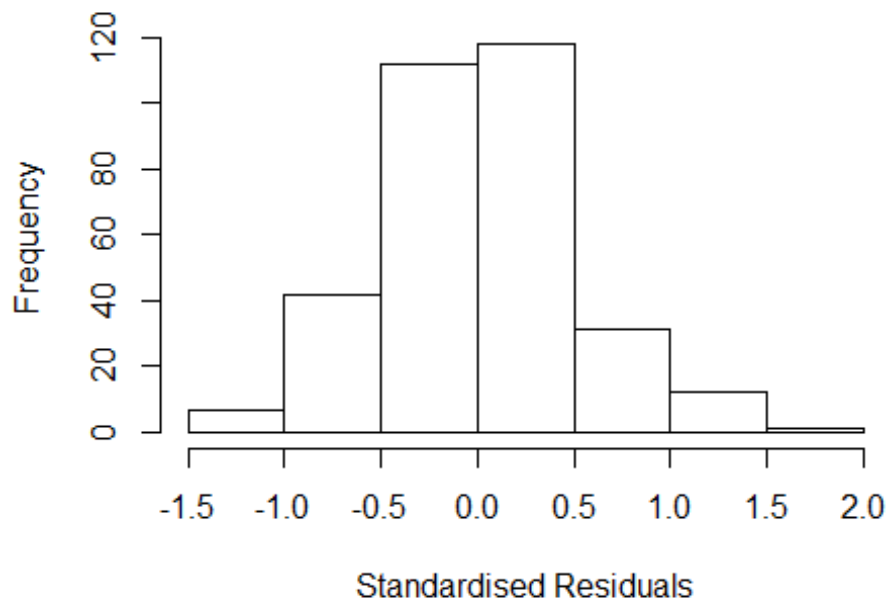
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.04416725424352"
## [1] "Male last author team size 2018 geometric mean: 2.33376074823249"

## 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 = 0.5
## 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 1.757 1      1.326
## LastAuthorFemale  1.961 1      1.400
## UniqueAuthors    2.983 4      1.146
## Year              3.954 16     1.044
```

Residuals from first and last author and team size



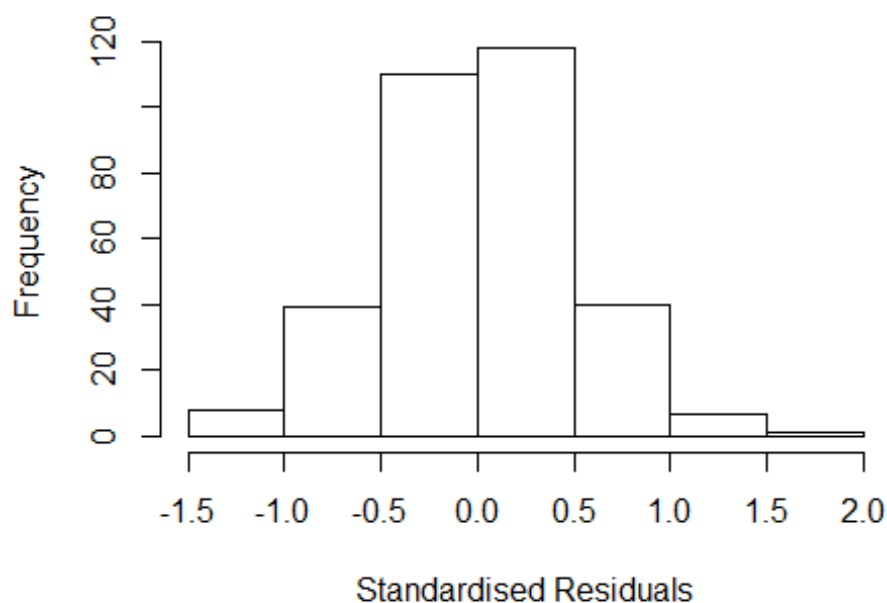
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.25115 -0.30688 0.00816 0.32115 1.82460
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0582 0.2262 4.68 4.4e-06 ***
## FirstAuthorFemale1 -0.0739 0.0702 -1.05 0.2934
## LastAuthorFemale1 -0.0429 0.0736 -0.58 0.5601
## UniqueAuthors2 0.1268 0.0684 1.86 0.0645 .
## UniqueAuthors3 0.2588 0.0835 3.10 0.0021 **
## UniqueAuthors4 0.1069 0.1368 0.78 0.4354
## UniqueAuthors5 0.1225 0.1093 1.12 0.2635
## Year1997 0.0593 0.2929 0.20 0.8398
## Year1998 -0.0710 0.2687 -0.26 0.7917
## Year1999 -0.3413 0.2728 -1.25 0.2120
```

```

## Year2000          0.2030      0.3292      0.62      0.5380
## Year2001         -0.1010      0.2515     -0.40      0.6882
## Year2002          0.1420      0.3137      0.45      0.6511
## Year2003          0.1317      0.2489      0.53      0.5971
## Year2004         -0.0809      0.2551     -0.32      0.7515
## Year2005         -0.2114      0.2564     -0.82      0.4103
## Year2006          0.3172      0.2468      1.29      0.1996
## Year2007          0.2584      0.2485      1.04      0.2992
## Year2008          0.0265      0.2656      0.10      0.9206
## Year2009         -0.0266      0.2469     -0.11      0.9144
## Year2010          0.1168      0.2430      0.48      0.6311
## Year2011         -0.1545      0.2463     -0.63      0.5310
## Year2012          0.0135      0.2462      0.05      0.9563
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.474
## Multiple R-squared:  0.119, Adjusted R-squared:  0.0544
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 303 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.106  0.877  0.956  0.903  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      3.10e-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 1.659 1      1.288
## LastAuthorFemale  1.878 1      1.370
## Year              1.568 16      1.014

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.3063 -0.3142 0.0187 0.2885 1.7558
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.08093 0.24371 4.44 1.3e-05 ***
## FirstAuthorFemale1 -0.06208 0.06997 -0.89 0.38
## LastAuthorFemale1 -0.05091 0.07381 -0.69 0.49
## Year1997 0.15451 0.32182 0.48 0.63
## Year1998 -0.03887 0.27885 -0.14 0.89
## Year1999 -0.25782 0.27064 -0.95 0.34
## Year2000 0.25813 0.33905 0.76 0.45
## Year2001 -0.08060 0.26684 -0.30 0.76
## Year2002 0.20616 0.32356 0.64 0.52
## Year2003 0.14272 0.26571 0.54 0.59
## Year2004 0.00138 0.26415 0.01 1.00
## Year2005 -0.13761 0.27464 -0.50 0.62
```

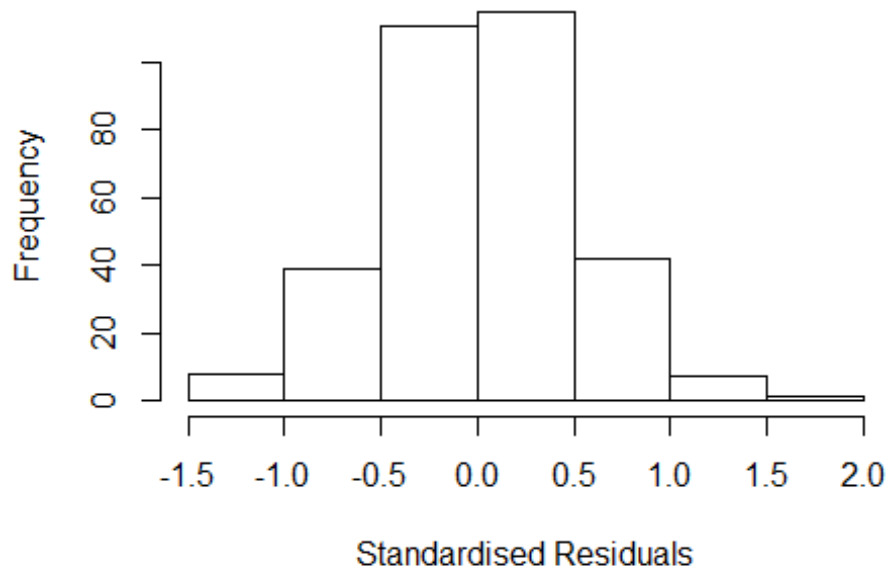


```

## Year2006          0.36327    0.26292    1.38    0.17
## Year2007          0.33835    0.26511    1.28    0.20
## Year2008          0.12573    0.28607    0.44    0.66
## Year2009          0.06264    0.26155    0.24    0.81
## Year2010          0.18541    0.25854    0.72    0.47
## Year2011         -0.09228    0.26124   -0.35    0.72
## Year2012          0.07572    0.26426    0.29    0.77
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.478
## Multiple R-squared:  0.091, Adjusted R-squared:  0.0372
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 297 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.149  0.868  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      3.10e-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 1.196 1      1.093
## Year              1.196 16      1.006

```

Residuals from first author

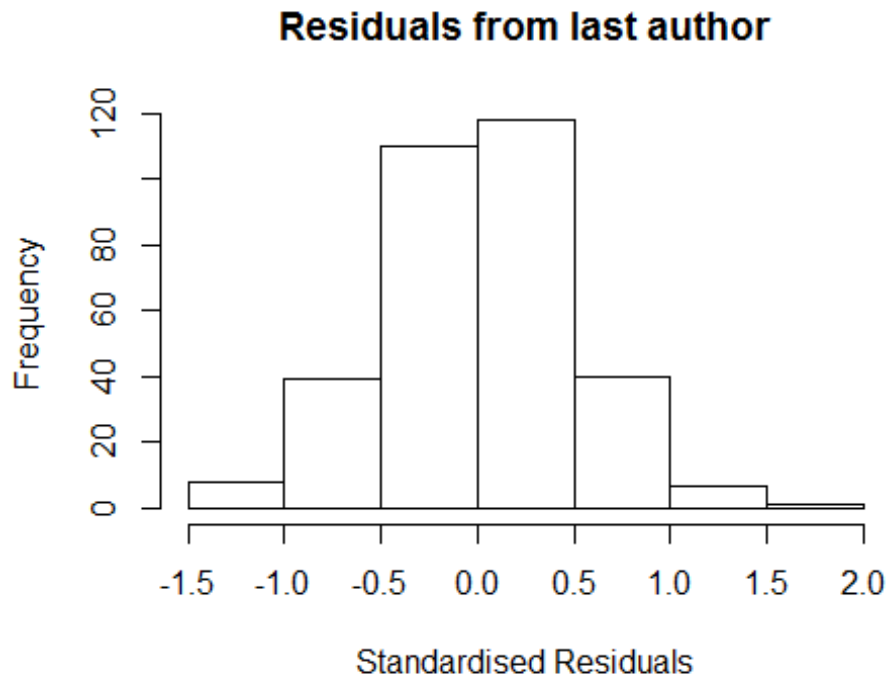


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3198 -0.3282 0.0206 0.2851 1.7727
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.07351 0.24217 4.43 1.3e-05 ***
## FirstAuthorFemale1 -0.08991 0.05941 -1.51 0.13
## Year1997 0.15227 0.32216 0.47 0.64
## Year1998 -0.04304 0.27565 -0.16 0.88
## Year1999 -0.24423 0.26902 -0.91 0.36
## Year2000 0.25305 0.34235 0.74 0.46
## Year2001 -0.08893 0.26361 -0.34 0.74
## Year2002 0.22096 0.31617 0.70 0.49
## Year2003 0.14092 0.26356 0.53 0.59
## Year2004 -0.00927 0.26069 -0.04 0.97
## Year2005 -0.13542 0.27274 -0.50 0.62
## Year2006 0.35380 0.25946 1.36 0.17
```

```

## Year2007          0.33621    0.26340    1.28    0.20
## Year2008          0.11540    0.28160    0.41    0.68
## Year2009          0.05889    0.25874    0.23    0.82
## Year2010          0.17862    0.25556    0.70    0.49
## Year2011         -0.09425    0.25879   -0.36    0.72
## Year2012          0.06572    0.26006    0.25    0.80
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.489
## Multiple R-squared:  0.0883, Adjusted R-squared:  0.0374
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 298 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.161  0.871  0.956  0.904  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      3.10e-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 1.362 1          1.167
## Year            1.362 16          1.010

```



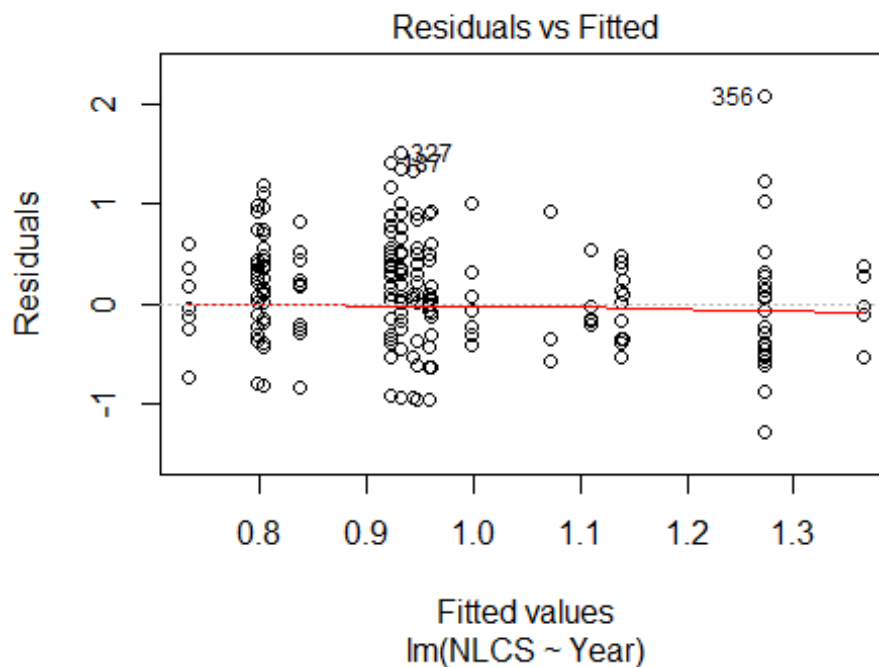
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3176 -0.3129 0.0221 0.2940 1.7678
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.06174 0.23636 4.49 1e-05 ***
## LastAuthorFemale1 -0.08440 0.06276 -1.34 0.18
## Year1997 0.15941 0.31733 0.50 0.62
## Year1998 -0.03493 0.27220 -0.13 0.90
## Year1999 -0.25287 0.26493 -0.95 0.34
## Year2000 0.26303 0.32591 0.81 0.42
## Year2001 -0.06768 0.26120 -0.26 0.80
## Year2002 0.18985 0.31309 0.61 0.54
## Year2003 0.15221 0.25865 0.59 0.56
## Year2004 0.00917 0.25858 0.04 0.97
## Year2005 -0.13519 0.26804 -0.50 0.61
## Year2006 0.37051 0.25601 1.45 0.15
```

```

## Year2007      0.34022      0.25904      1.31      0.19
## Year2008      0.14315      0.27985      0.51      0.61
## Year2009      0.06160      0.25460      0.24      0.81
## Year2010      0.18698      0.25187      0.74      0.46
## Year2011     -0.07868      0.25444     -0.31      0.76
## Year2012      0.07964      0.25760      0.31      0.76
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared:  0.0886, Adjusted R-squared:  0.0378
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 297 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.145  0.869  0.953  0.900  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      3.10e-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: 323"
## [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
##    7    5   13    9    7    7   10   13    4   11   10   20   50   44   32
## 2011 2012
##   41   35
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    3   11    8    5    5    9   11    3    7   10   17   43   38   25
## 2011 2012

```

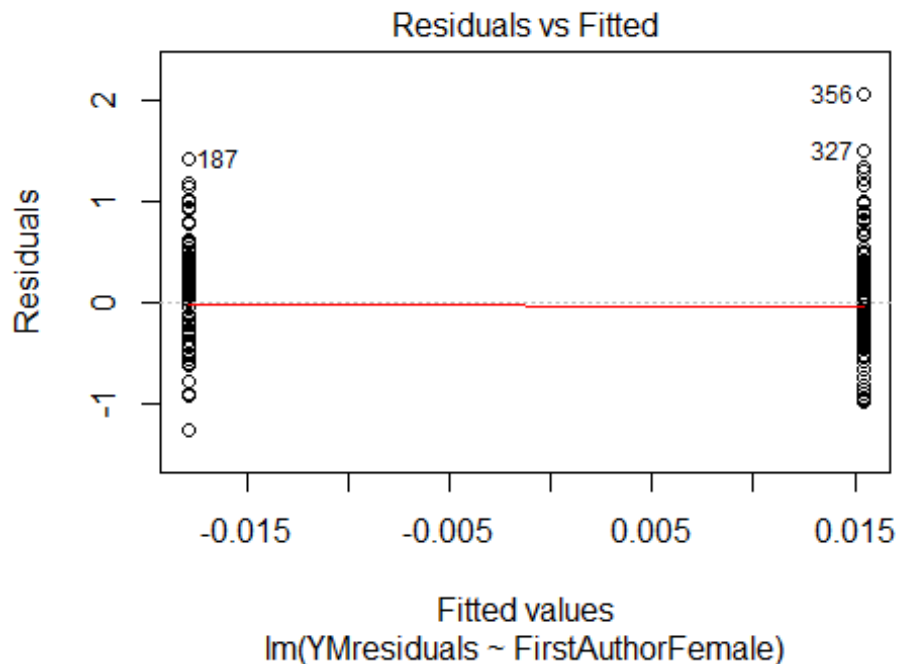
```
## 33 24
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 2 11 8 4 5 9 11 3 7 10 16 41 34 24
## 2011 2012
## 31 23
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS 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 = 0.47, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 1.94653971032285"
## [1] "Male first author team size 2018 geometric mean: 1.89923919456909"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

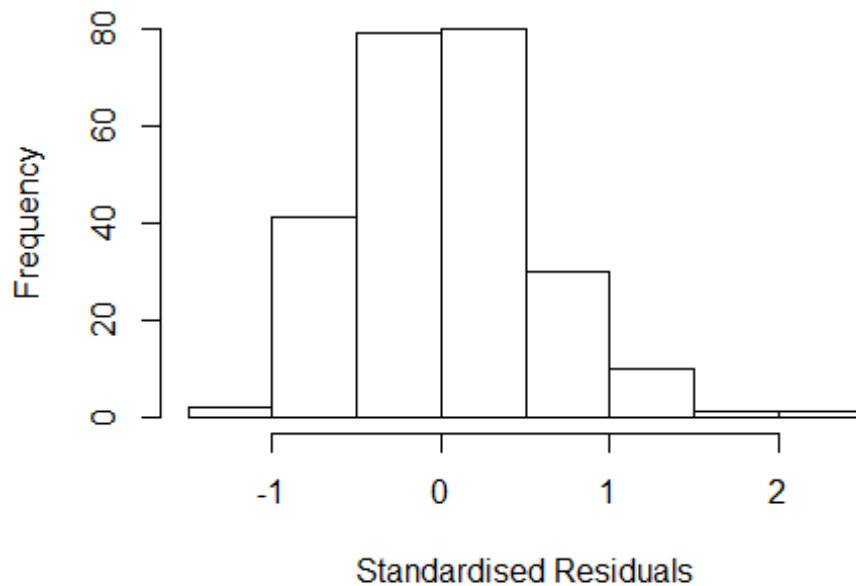
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.87178239219018"
## [1] "Male last author team size 2018 geometric mean: 1.96242858209132"

## 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 = 0.8
## 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 2.571  1      1.603
## LastAuthorFemale  2.746  1      1.657
## UniqueAuthors    2.525  4      1.123
## Year              3.731 16      1.042
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.1380 -0.3504 0.0088 0.3864 2.1959
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.10082 0.13057 8.43 4.4e-15 ***
## FirstAuthorFemale1 0.00817 0.10980 0.07 0.9408
## LastAuthorFemale1 -0.02526 0.10971 -0.23 0.8181
## UniqueAuthors2 0.01949 0.09767 0.20 0.8420
## UniqueAuthors3 0.16835 0.10993 1.53 0.1271
## UniqueAuthors4 0.24459 0.24462 1.00 0.3185
## UniqueAuthors5 0.21947 0.24875 0.88 0.3786
## Year1997 -0.48477 0.14996 -3.23 0.0014 **
## Year1998 -0.26286 0.20733 -1.27 0.2062
## Year1999 -0.15198 0.20448 -0.74 0.4581
```

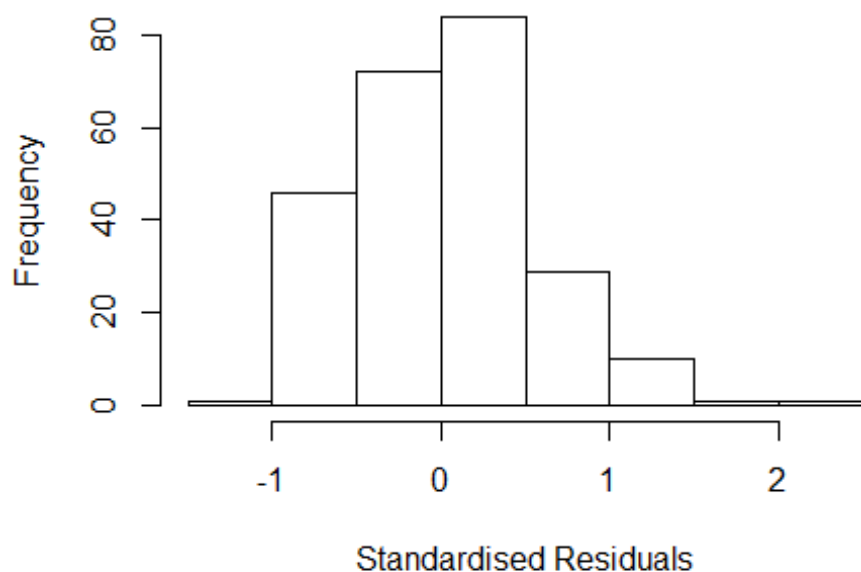


```

## Year2000      -0.47624    0.26488   -1.80    0.0735 .
## Year2001      0.22800    0.19760    1.15    0.2498
## Year2002      0.01598    0.17744    0.09    0.9283
## Year2003     -0.15651    0.20950   -0.75    0.4558
## Year2004     -0.01791    0.17971   -0.10    0.9207
## Year2005     -0.38843    0.21717   -1.79    0.0750 .
## Year2006     -0.19123    0.20475   -0.93    0.3513
## Year2007     -0.12370    0.24268   -0.51    0.6107
## Year2008     -0.23978    0.16090   -1.49    0.1376
## Year2009     -0.31392    0.16093   -1.95    0.0524 .
## Year2010     -0.33440    0.18253   -1.83    0.0683 .
## Year2011     -0.26934    0.19389   -1.39    0.1662
## Year2012      0.05425    0.19891    0.27    0.7853
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.0759, Adjusted R-squared:  -0.0161
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 224 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.157  0.858  0.961  0.917  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      4.10e-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 2.634 1      1.623
## LastAuthorFemale 2.887 1      1.699
## Year      1.721 16      1.017

```

Residuals from first and last author



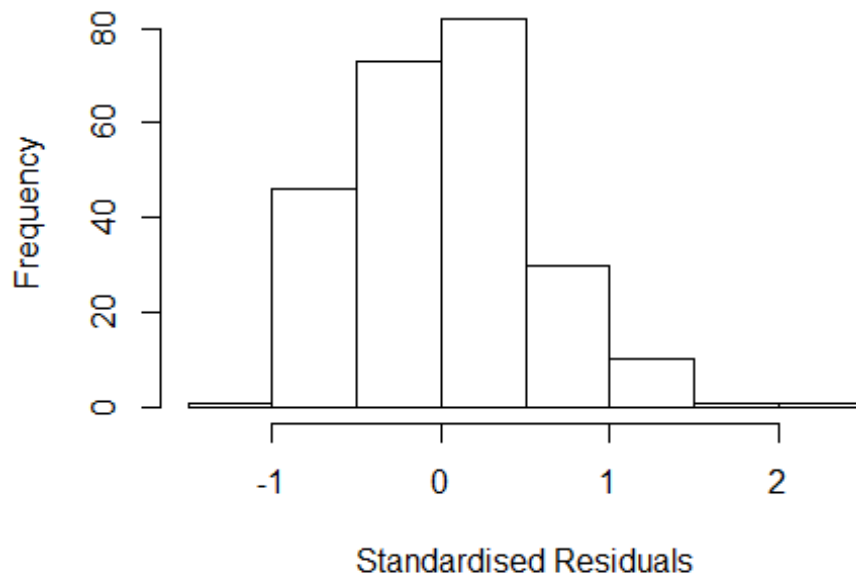
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1913 -0.3766 0.0175 0.3892 2.1518
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1030 0.1272 8.67 8.6e-16 ***
## FirstAuthorFemale1 0.0127 0.1138 0.11 0.9112
## LastAuthorFemale1 -0.0206 0.1145 -0.18 0.8576
## Year1997 -0.4916 0.1497 -3.28 0.0012 **
## Year1998 -0.2490 0.2099 -1.19 0.2368
## Year1999 -0.1315 0.1970 -0.67 0.5051
## Year2000 -0.4811 0.2640 -1.82 0.0698 .
## Year2001 0.2673 0.1910 1.40 0.1629
## Year2002 0.0358 0.1769 0.20 0.8398
## Year2003 -0.1344 0.2054 -0.65 0.5134
## Year2004 0.0351 0.1947 0.18 0.8570
## Year2005 -0.3585 0.2048 -1.75 0.0813 .
```

```

## Year2006          -0.1527      0.1952   -0.78   0.4349
## Year2007          -0.1259      0.2428   -0.52   0.6047
## Year2008          -0.2008      0.1599   -1.26   0.2105
## Year2009          -0.2974      0.1622   -1.83   0.0680 .
## Year2010          -0.3243      0.1835   -1.77   0.0785 .
## Year2011          -0.2297      0.1957   -1.17   0.2418
## Year2012           0.0962      0.1842    0.52   0.6022
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.605
## Multiple R-squared:  0.0645, Adjusted R-squared:  -0.0104
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 226 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.180  0.858  0.961  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      4.10e-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 1.286 1      1.134
## Year              1.286 16      1.008

```

Residuals from first author



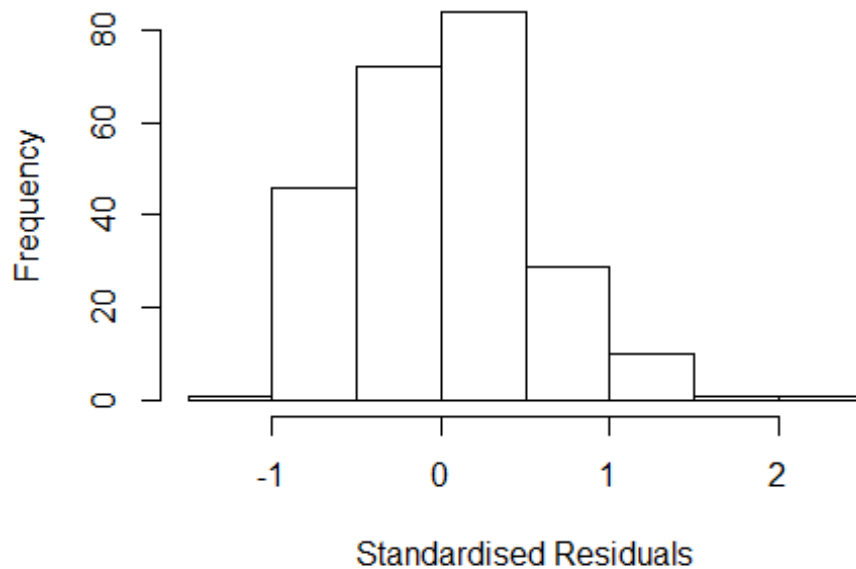
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1949 -0.3774 0.0165 0.3965 2.1540
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1018 0.1273 8.65 9.6e-16 ***
## FirstAuthorFemale1 -0.0021 0.0795 -0.03 0.9790
## Year1997 -0.4932 0.1509 -3.27 0.0013 **
## Year1998 -0.2530 0.2069 -1.22 0.2227
## Year1999 -0.1322 0.1970 -0.67 0.5030
## Year2000 -0.4814 0.2647 -1.82 0.0703 .
## Year2001 0.2676 0.1919 1.39 0.1646
## Year2002 0.0376 0.1767 0.21 0.8316
## Year2003 -0.1362 0.2050 -0.66 0.5072
## Year2004 0.0413 0.1939 0.21 0.8316
## Year2005 -0.3630 0.2021 -1.80 0.0737 .
## Year2006 -0.1522 0.1951 -0.78 0.4361
```

```

## Year2007          -0.1307      0.2392   -0.55   0.5853
## Year2008          -0.2049      0.1580   -1.30   0.1959
## Year2009          -0.2977      0.1625   -1.83   0.0682 .
## Year2010          -0.3259      0.1835   -1.78   0.0772 .
## Year2011          -0.2285      0.1965   -1.16   0.2461
## Year2012           0.0952      0.1846    0.52   0.6067
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.0644, Adjusted R-squared:  -0.006
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 226 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.177  0.856  0.961  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      4.10e-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 1.416 1          1.190
## Year              1.416 16          1.011

```

Residuals from last author



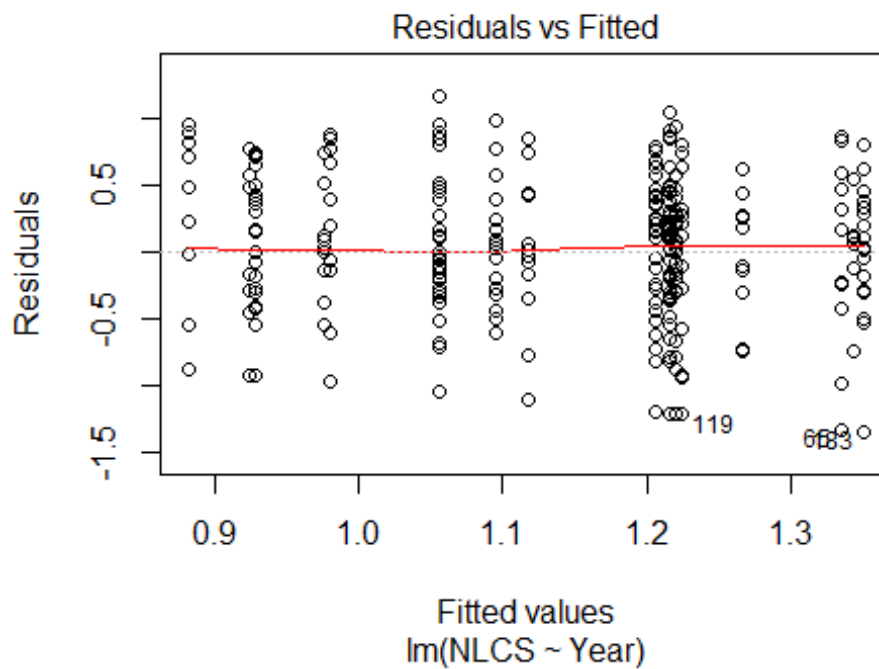
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1918 -0.3777 0.0149 0.3915 2.1474
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1040 0.1268 8.71 6.6e-16 ***
## LastAuthorFemale1 -0.0118 0.0803 -0.15 0.8836
## Year1997 -0.4906 0.1486 -3.30 0.0011 **
## Year1998 -0.2512 0.2082 -1.21 0.2288
## Year1999 -0.1305 0.1973 -0.66 0.5090
## Year2000 -0.4812 0.2631 -1.83 0.0687 .
## Year2001 0.2669 0.1903 1.40 0.1621
## Year2002 0.0371 0.1761 0.21 0.8331
## Year2003 -0.1336 0.2046 -0.65 0.5143
## Year2004 0.0383 0.1940 0.20 0.8437
## Year2005 -0.3594 0.2048 -1.76 0.0806 .
## Year2006 -0.1501 0.1941 -0.77 0.4400
```

```

## Year2007          -0.1276      0.2397   -0.53    0.5951
## Year2008          -0.1998      0.1599   -1.25    0.2128
## Year2009          -0.2950      0.1607   -1.84    0.0677 .
## Year2010          -0.3234      0.1830   -1.77    0.0785 .
## Year2011          -0.2263      0.1950   -1.16    0.2471
## Year2012           0.0996      0.1838    0.54    0.5886
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.613
## Multiple R-squared:  0.0643, Adjusted R-squared:  -0.00604
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 17 weights are ~= 1. The remaining 227 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.195  0.862  0.963   0.920  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      4.10e-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: 244"
## [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
##   12  13   7   8  14  13  14  11  18  21  21  17  15  35  43
## 2011 2012
##   33  47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12  12   7   7  11  11  12   9  16  18  17  16  12  32  40
## 2011 2012

```

```
## 31 42
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 12 7 6 11 9 12 8 14 16 16 14 10 28 36
## 2011 2012
## 30 38
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.7
```

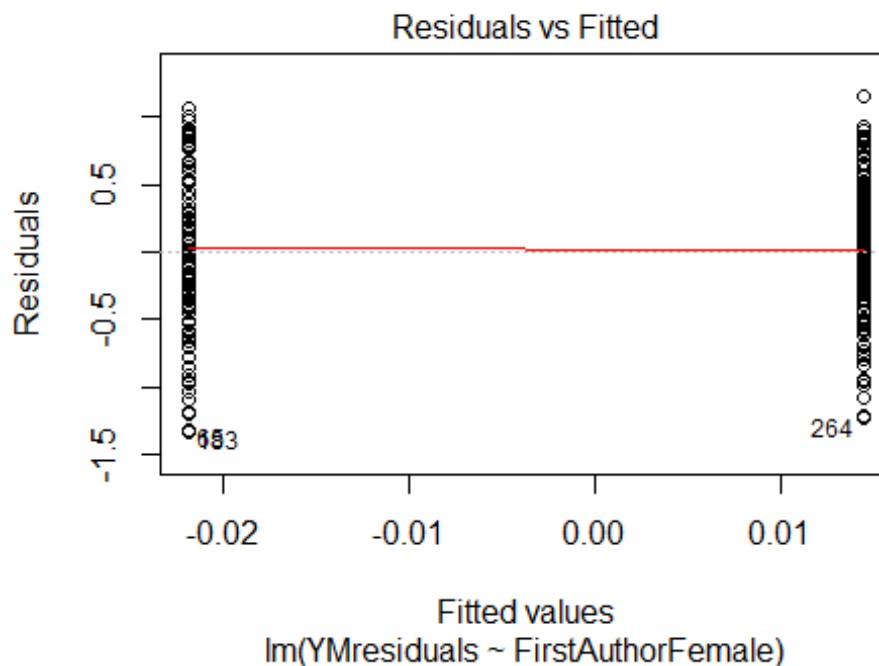


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9.3, df = 1, p-value = 0.002
## [1] "Female first author team size 2018 geometric mean: 1.72788704583333"
## [1] "Male first author team size 2018 geometric mean: 1.84218488139299"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



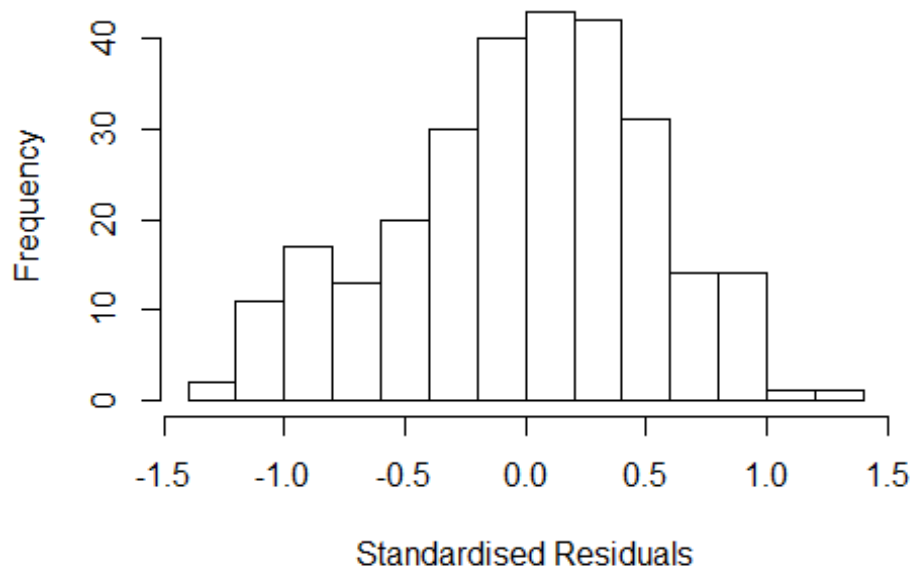
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150, p-value = 0.8
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.83523578992857"
## [1] "Male last author team size 2018 geometric mean: 1.66851044102683"

## 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 = 0.8
## 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 1.742  1      1.320
## LastAuthorFemale  2.137  1      1.462
## UniqueAuthors    2.690  4      1.132
## Year              3.188 16      1.037
```

Residuals from first and last author and team size



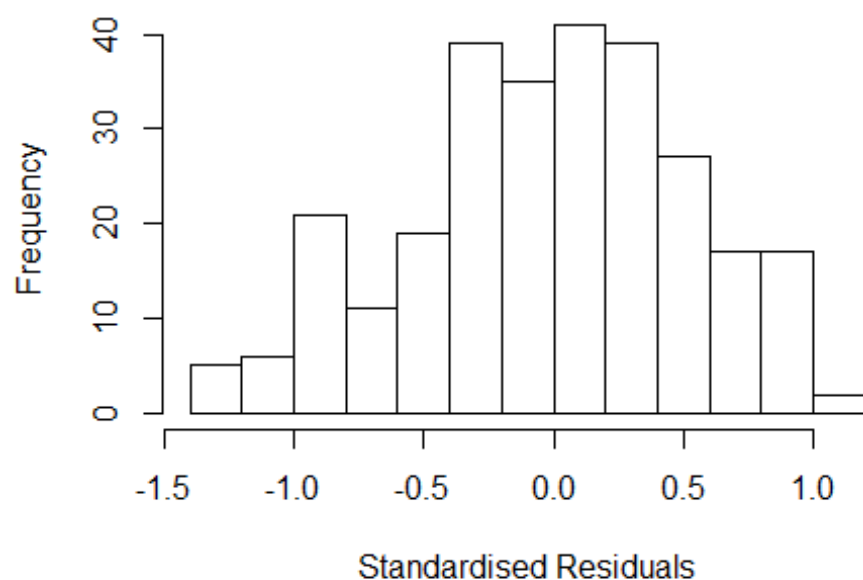
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2829 -0.3252 0.0291 0.3650 1.2147
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8718 0.3243 2.69 0.0077 **
## FirstAuthorFemale1 -0.1423 0.0886 -1.61 0.1094
## LastAuthorFemale1 0.2397 0.0890 2.69 0.0075 **
## UniqueAuthors2 0.1488 0.0773 1.92 0.0554 .
## UniqueAuthors3 0.4125 0.1530 2.70 0.0075 **
## UniqueAuthors4 0.1625 0.2002 0.81 0.4180
## UniqueAuthors5 0.3560 0.2225 1.60 0.1109
## Year1997 0.0192 0.4146 0.05 0.9631
## Year1998 0.0219 0.4279 0.05 0.9592
## Year1999 0.4012 0.3633 1.10 0.2704
```

```

## Year2000          0.4111      0.3859      1.07      0.2877
## Year2001          0.2560      0.3531      0.72      0.4692
## Year2002          0.0385      0.3418      0.11      0.9105
## Year2003         -0.1083      0.3610     -0.30      0.7645
## Year2004          0.1452      0.3669      0.40      0.6927
## Year2005          0.0948      0.3375      0.28      0.7790
## Year2006          0.3221      0.3551      0.91      0.3652
## Year2007         -0.1999      0.3547     -0.56      0.5735
## Year2008         -0.0306      0.3623     -0.08      0.9328
## Year2009          0.2849      0.3356      0.85      0.3968
## Year2010          0.0381      0.3320      0.11      0.9087
## Year2011          0.2268      0.3367      0.67      0.5012
## Year2012          0.1787      0.3360      0.53      0.5952
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.516
## Multiple R-squared:  0.137, Adjusted R-squared:  0.063
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.517  0.868  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      3.58e-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 1.845 1      1.358
## LastAuthorFemale  2.084 1      1.443
## Year              1.404 16      1.011

```

Residuals from first and last author



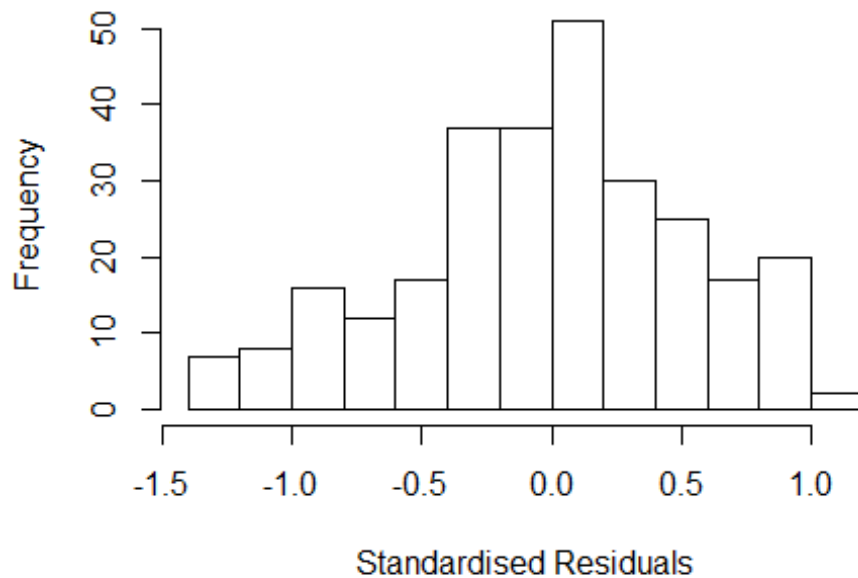
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.359 -0.336 0.019 0.346 1.146
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8697 0.3096 2.81 0.0053 **
## FirstAuthorFemale1 -0.1172 0.0956 -1.23 0.2214
## LastAuthorFemale1 0.2193 0.0950 2.31 0.0218 *
## Year1997 0.0903 0.4101 0.22 0.8259
## Year1998 0.0218 0.4046 0.05 0.9571
## Year1999 0.4544 0.3494 1.30 0.1946
## Year2000 0.4896 0.3843 1.27 0.2038
## Year2001 0.2703 0.3386 0.80 0.4254
## Year2002 0.0592 0.3293 0.18 0.8573
## Year2003 -0.0254 0.3395 -0.07 0.9403
## Year2004 0.2406 0.3623 0.66 0.5072
## Year2005 0.1768 0.3247 0.54 0.5865
```

```

## Year2006          0.4214      0.3329      1.27      0.2068
## Year2007         -0.0766      0.3528     -0.22      0.8282
## Year2008          0.0495      0.3481      0.14      0.8870
## Year2009          0.3761      0.3189      1.18      0.2393
## Year2010          0.1041      0.3199      0.33      0.7452
## Year2011          0.3170      0.3227      0.98      0.3267
## Year2012          0.2731      0.3189      0.86      0.3927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.538
## Multiple R-squared:  0.0948, Adjusted R-squared:  0.0321
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 20 weights are ~= 1. The remaining 259 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.503  0.846  0.954  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      3.58e-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 1.199 1      1.095
## Year              1.199 16      1.006

```

Residuals from first author



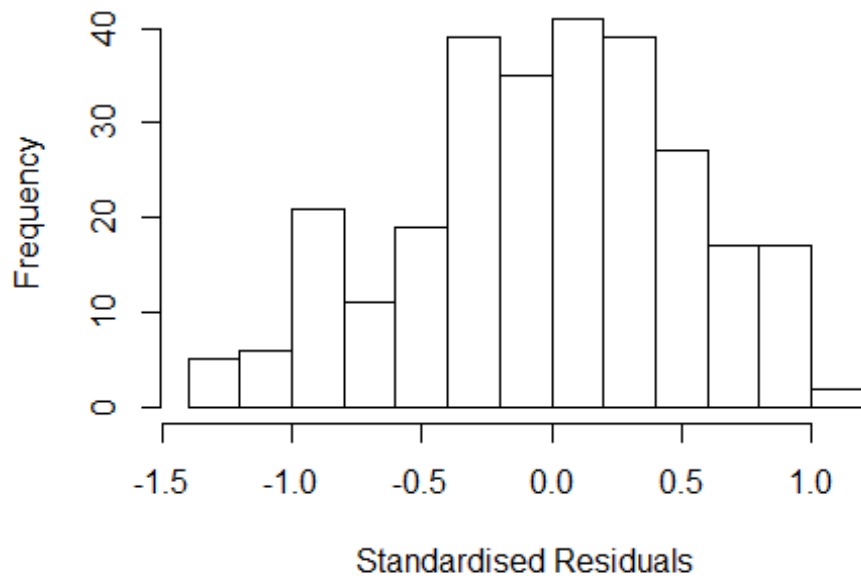
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3907 -0.3374 0.0242 0.3554 1.1735
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8879 0.3187 2.79 0.0057 **
## FirstAuthorFemale1 0.0298 0.0779 0.38 0.7030
## Year1997 0.1151 0.4210 0.27 0.7848
## Year1998 0.0320 0.4104 0.08 0.9378
## Year1999 0.4395 0.3527 1.25 0.2139
## Year2000 0.5027 0.3980 1.26 0.2077
## Year2001 0.2700 0.3472 0.78 0.4376
## Year2002 0.0638 0.3396 0.19 0.8511
## Year2003 -0.0211 0.3456 -0.06 0.9515
## Year2004 0.2821 0.3702 0.76 0.4468
## Year2005 0.1853 0.3373 0.55 0.5833
## Year2006 0.4566 0.3408 1.34 0.1815
```

```

## Year2007          -0.0697      0.3577   -0.19   0.8457
## Year2008          0.0983      0.3627    0.27   0.7865
## Year2009          0.3948      0.3285    1.20   0.2304
## Year2010          0.1308      0.3316    0.39   0.6935
## Year2011          0.3552      0.3323    1.07   0.2861
## Year2012          0.3041      0.3288    0.93   0.3558
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.534
## Multiple R-squared:  0.0788, Adjusted R-squared:  0.0188
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.477  0.840   0.954   0.903   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      3.58e-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 1.288 1          1.135
## Year            1.288 16          1.008

```

Residuals from last author



```
## [1] "List of 0 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
## -1.3588 -0.3442 0.0302 0.3339 1.1338
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8596 0.3033 2.83 0.005 **
## LastAuthorFemale1 0.1417 0.0746 1.90 0.059 .
## Year1997 0.0766 0.3993 0.19 0.848
## Year1998 0.0162 0.4002 0.04 0.968
## Year1999 0.4251 0.3423 1.24 0.215
## Year2000 0.4723 0.3725 1.27 0.206
## Year2001 0.2447 0.3343 0.73 0.465
## Year2002 0.0389 0.3255 0.12 0.905
## Year2003 -0.0548 0.3308 -0.17 0.868
## Year2004 0.2291 0.3584 0.64 0.523
## Year2005 0.1575 0.3189 0.49 0.622
## Year2006 0.4026 0.3247 1.24 0.216
```

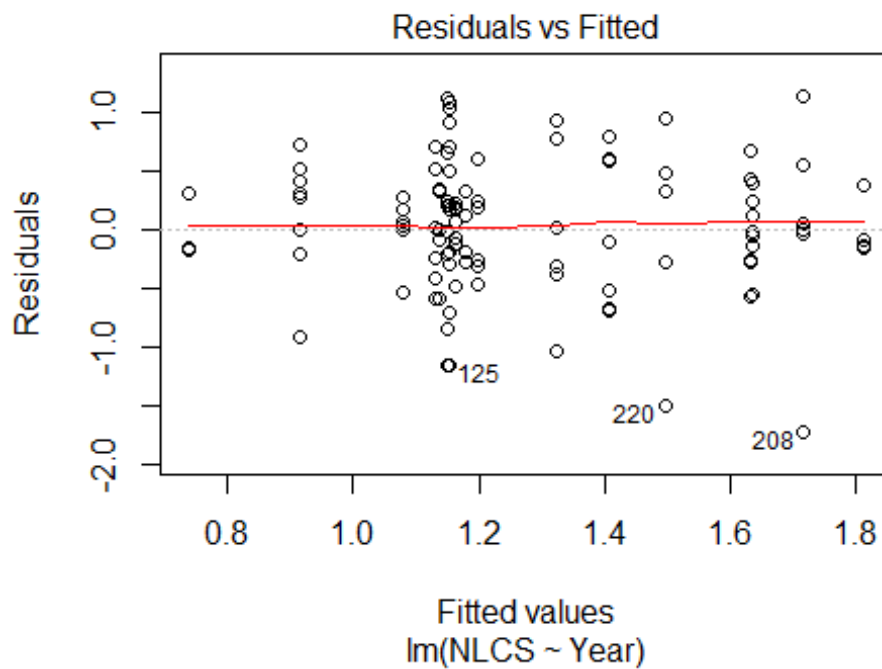


```

## Year2007          -0.0883      0.3464    -0.25      0.799
## Year2008           0.0319      0.3426      0.09      0.926
## Year2009           0.3575      0.3107      1.15      0.251
## Year2010           0.0868      0.3155      0.28      0.783
## Year2011           0.2977      0.3141      0.95      0.344
## Year2012           0.2570      0.3117      0.82      0.410
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.543
## Multiple R-squared:  0.0899, Adjusted R-squared:  0.0306
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 260 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.510  0.852  0.954  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      3.58e-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: 279"
## [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
##   10   6   8   6   8  13   8  14   7   8  10  14   8   6   7
## 2011 2012
##    7    7
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    3    6    4    5    6    8  10    6    7    7  13    4    6    5
## 2011 2012

```

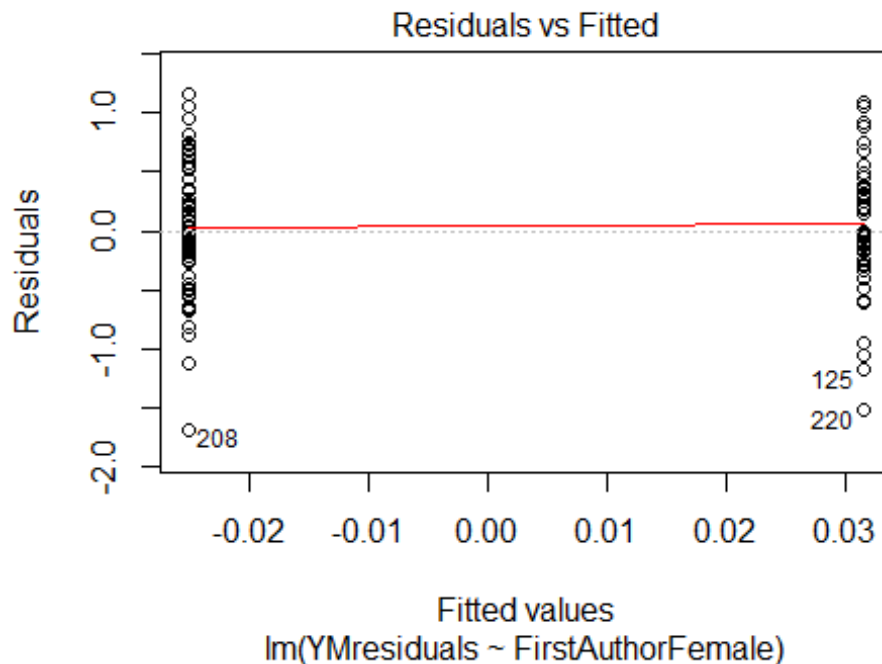
```
##      6      5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      7      3      6      4      5      5      7      10      6      7      7      13      4      6      4
## 2011 2012
##      5      5
## [1] "Heteroscedasticity checks, confirming that there are 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 = 0.017, df = 1, p-value = 0.9
## [1] "Female first author team size 2018 geometric mean: 3.46410161513775"
## [1] "Male first author team size 2018 geometric mean: 1.25992104989487"
## 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 = 11, p-value = 0.09
## 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: 1.57461010625845"

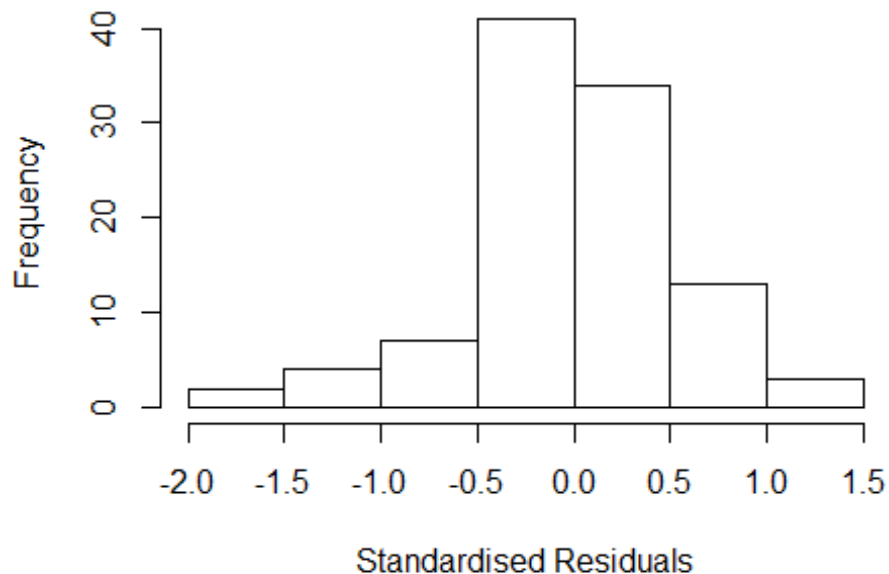
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5, 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	3.588	1	1.894
LastAuthorFemale	6.891	1	2.625
UniqueAuthors	51.776	4	1.638
Year	248.224	16	1.188

Residuals from first and last author and team size



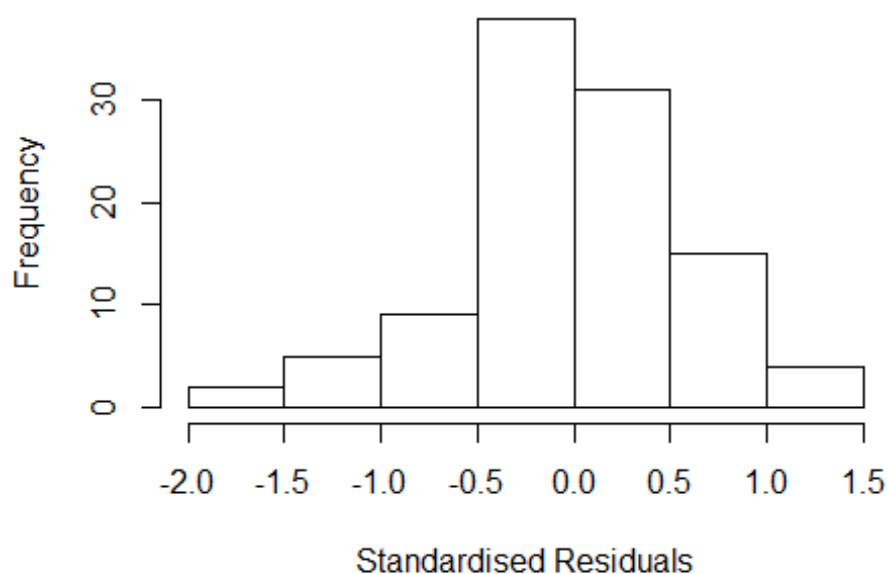
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.9098 -0.2894 -0.0474 0.3229 1.2737
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9096 0.1471 6.18 2.4e-08 ***
## FirstAuthorFemale1 0.0779 0.1397 0.56 0.57863
## LastAuthorFemale1 0.1216 0.2197 0.55 0.58122
## UniqueAuthors2 0.4462 0.1160 3.85 0.00024 ***
## UniqueAuthors3 0.2374 0.1848 1.28 0.20276
## UniqueAuthors4 0.4025 0.3043 1.32 0.18966
## UniqueAuthors5 0.1118 0.2698 0.41 0.67985
## Year1997 -0.3201 0.2712 -1.18 0.24140
## Year1998 0.1166 0.2617 0.45 0.65706
## Year1999 0.2189 0.1560 1.40 0.16427
```

```

## Year2000          -0.0806      0.1616    -0.50   0.61935
## Year2001          -0.0921      0.2068    -0.45   0.65736
## Year2002          -0.1028      0.3101    -0.33   0.74100
## Year2003          -0.0861      0.2268    -0.38   0.70520
## Year2004           0.1168      0.3823     0.31   0.76070
## Year2005           0.3030      0.2678     1.13   0.26123
## Year2006           0.4943      0.1676     2.95   0.00416 **
## Year2007           0.0861      0.4377     0.20   0.84464
## Year2008           0.3683      0.1679     2.19   0.03118 *
## Year2009           0.0105      0.1640     0.06   0.94888
## Year2010           0.2974      0.2441     1.22   0.22669
## Year2011           1.0003      0.3812     2.62   0.01039 *
## Year2012           0.6034      0.4947     1.22   0.22612
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.533
## Multiple R-squared:  0.302, Adjusted R-squared:  0.112
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 97 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.172  0.906  0.967  0.906  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      9.62e-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 2.467 1      1.571
## LastAuthorFemale  4.390 1      2.095
## Year              7.006 16      1.063

```

Residuals from first and last author



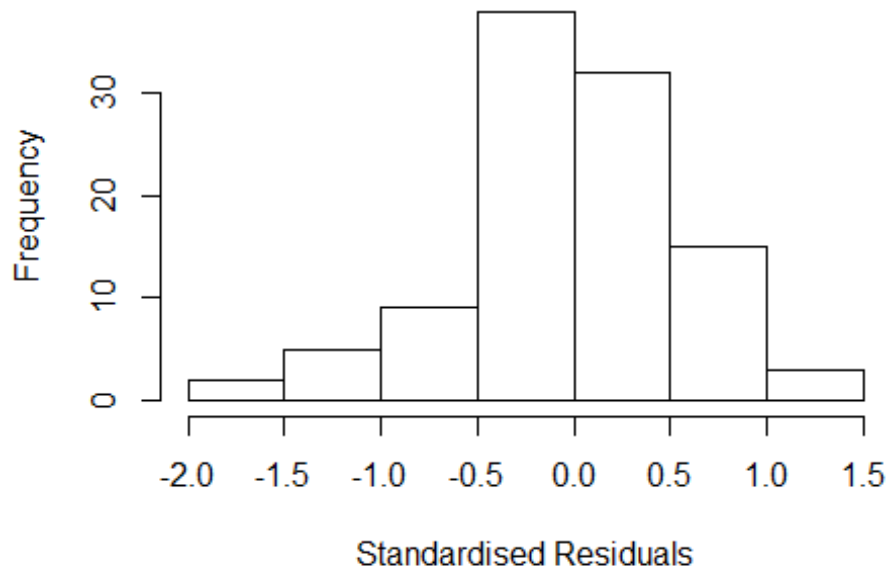
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.981 -0.304 -0.018 0.304 1.122
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1638 0.1384 8.41 8.4e-13 ***
## FirstAuthorFemale1 0.1939 0.1343 1.44 0.15267
## LastAuthorFemale1 -0.0918 0.1919 -0.48 0.63353
## Year1997 -0.4279 0.1887 -2.27 0.02588 *
## Year1998 -0.0437 0.2506 -0.17 0.86203
## Year1999 -0.0112 0.1573 -0.07 0.94349
## Year2000 -0.1209 0.1993 -0.61 0.54580
## Year2001 -0.1850 0.1887 -0.98 0.32975
## Year2002 -0.1080 0.3585 -0.30 0.76404
## Year2003 -0.2831 0.2152 -1.32 0.19182
## Year2004 0.0667 0.3884 0.17 0.86398
## Year2005 0.2200 0.2919 0.75 0.45320
```

```

## Year2006          0.4193      0.1535      2.73  0.00766 **
## Year2007         -0.0516      0.3708     -0.14  0.88964
## Year2008          0.5676      0.1508      3.76  0.00031 ***
## Year2009         -0.0419      0.2132     -0.20  0.84462
## Year2010          0.2415      0.2022      1.19  0.23555
## Year2011          0.8171      0.4096      1.99  0.04929 *
## Year2012          0.4543      0.4388      1.04  0.30337
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.584
## Multiple R-squared:  0.216, Adjusted R-squared:  0.0497
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 96 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.226  0.885  0.969  0.913  0.992  0.999
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol      solve.tol      eps.outlier      eps.x
##      1.00e-07      1.00e-07      9.62e-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 1.991 1          1.411
## Year              1.991 16          1.022

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.9913 -0.3131 -0.0357 0.3138 1.0847
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.12804 0.09601 11.75 < 2e-16 ***
## FirstAuthorFemale1 0.13530 0.13100 1.03 0.30457
## Year1997 -0.39215 0.16085 -2.44 0.01683 *
## Year1998 -0.00826 0.23290 -0.04 0.97179
## Year1999 0.01643 0.13449 0.12 0.90303
## Year2000 -0.09264 0.20292 -0.46 0.64916
## Year2001 -0.13125 0.14854 -0.88 0.37935
## Year2002 -0.06804 0.33772 -0.20 0.84081
## Year2003 -0.26770 0.21375 -1.25 0.21382
## Year2004 0.10033 0.37561 0.27 0.79002
## Year2005 0.25024 0.28166 0.89 0.37677
## Year2006 0.43696 0.14797 2.95 0.00406 **
```

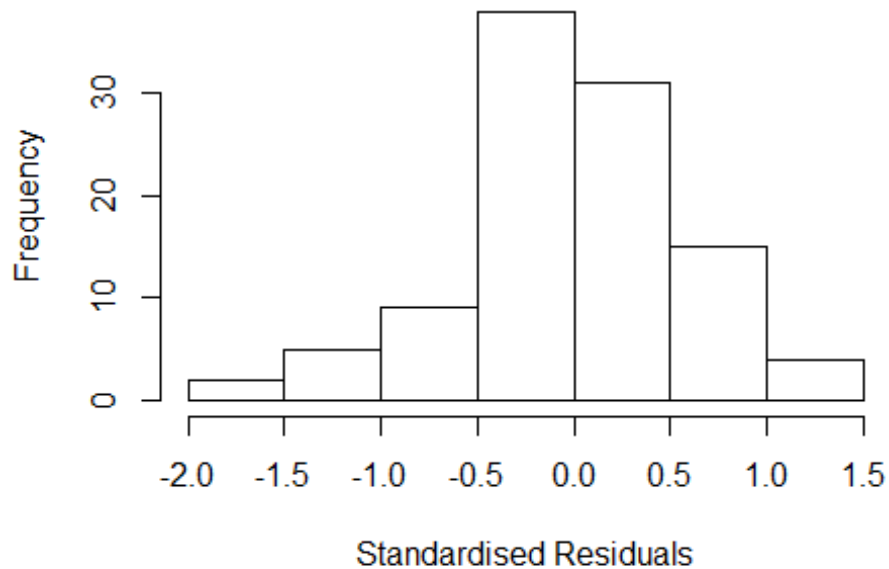


```

## Year2007          0.00300    0.31481    0.01  0.99242
## Year2008          0.57871    0.14703    3.94  0.00017 ***
## Year2009         -0.02974    0.21958   -0.14  0.89258
## Year2010          0.26056    0.18662    1.40  0.16625
## Year2011          0.86325    0.38814    2.22  0.02876 *
## Year2012          0.51903    0.36461    1.42  0.15820
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.575
## Multiple R-squared:  0.22,   Adjusted R-squared:  0.0653
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 97 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.205  0.882  0.968  0.911  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      9.62e-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 3.071 1          1.752
## Year            3.071 16          1.036

```

Residuals from last author



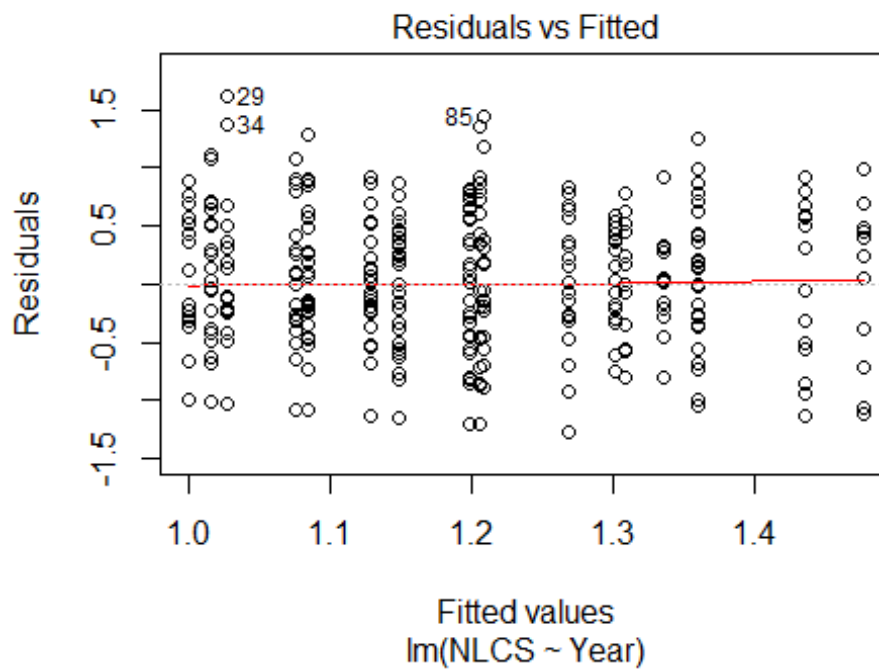
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.9921 -0.3246 -0.0333 0.2994 1.1165
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1374 0.1297 8.77 1.4e-13 ***
## LastAuthorFemale1 0.0519 0.1668 0.31 0.756
## Year1997 -0.4015 0.1825 -2.20 0.030 *
## Year1998 -0.0176 0.2466 -0.07 0.943
## Year1999 0.0275 0.1556 0.18 0.860
## Year2000 -0.0245 0.1853 -0.13 0.895
## Year2001 -0.0671 0.1630 -0.41 0.682
## Year2002 -0.0258 0.3551 -0.07 0.942
## Year2003 -0.2236 0.2026 -1.10 0.273
## Year2004 0.1583 0.3536 0.45 0.655
## Year2005 0.2537 0.2949 0.86 0.392
## Year2006 0.4737 0.1391 3.41 0.001 **
```

```

## Year2007          0.0523      0.3640      0.14      0.886
## Year2008          0.6312      0.1390      4.54      1.8e-05 ***
## Year2009          0.0192      0.1931      0.10      0.921
## Year2010          0.2911      0.1994      1.46      0.148
## Year2011          0.8547      0.4011      2.13      0.036 *
## Year2012          0.5566      0.4415      1.26      0.211
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.574
## Multiple R-squared:  0.21,   Adjusted R-squared:  0.0539
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 95 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.204  0.880  0.966  0.908  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      9.62e-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: 104"
## [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
##   14   23   29   30   21   39   38   26   21   18   15   35   25   35   29
## 2011 2012
##   50   25
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   11   19   22   19   14   17   30   20   18   14   12   30   21   27   23
## 2011 2012

```

```
## 34 23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 11 19 21 19 14 15 28 20 17 14 10 29 19 23 21
## 2011 2012
## 33 22
## [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
```



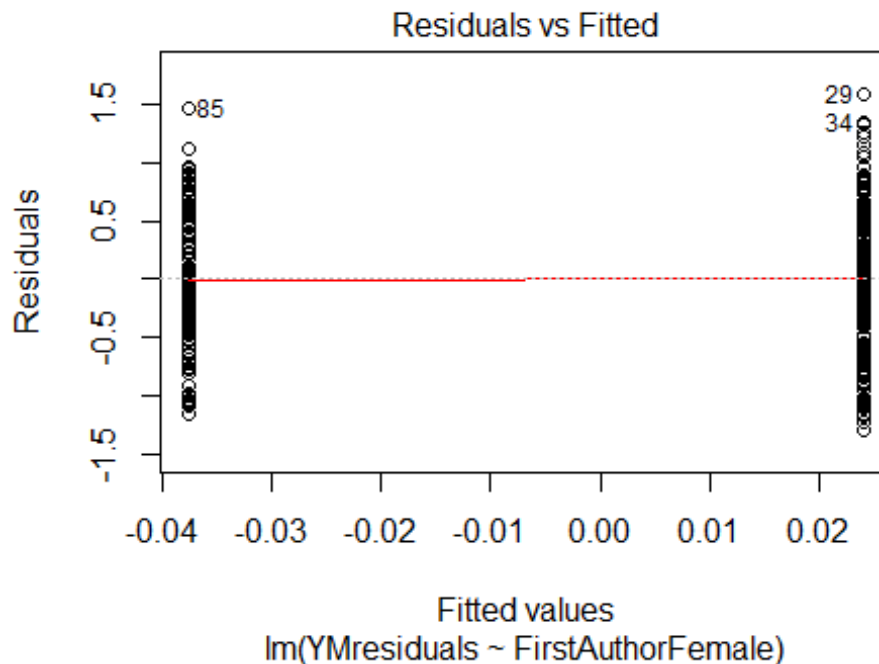
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.8, df = 1, p-value = 0.2

## [1] "Female first author team size 2018 geometric mean: 1.74549231584704"
## [1] "Male first author team size 2018 geometric mean: 1.84218488139299"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

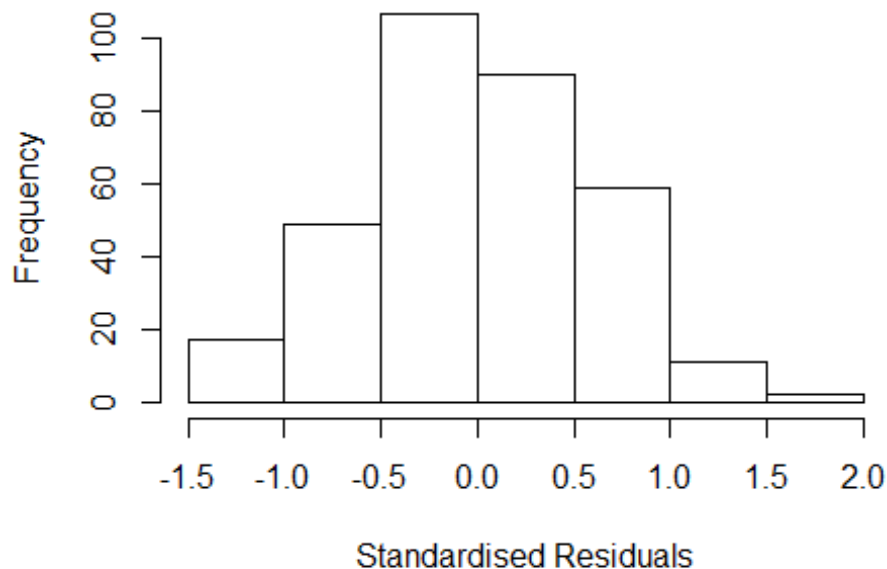
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 110, p-value = 0.7
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.73051460095418"
## [1] "Male last author team size 2018 geometric mean: 1.83773678485593"

## 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 = 0.8
## 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 1.818  1      1.348
## LastAuthorFemale  1.692  1      1.301
## UniqueAuthors    2.272  4      1.108
## Year              2.922 16      1.034
```

Residuals from first and last author and team size



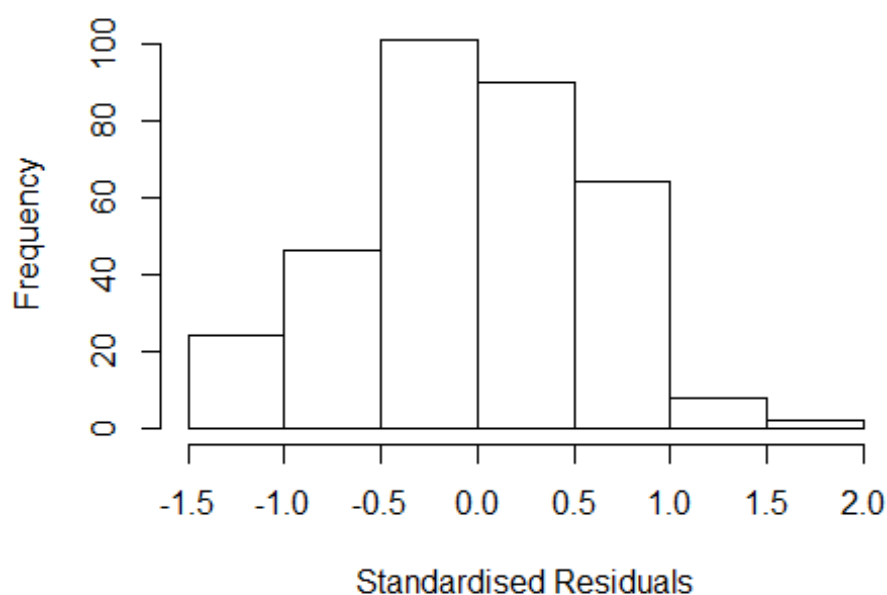
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.285 -0.373 -0.021 0.441 1.683
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4295 0.2251 6.35 7.5e-10 ***
## FirstAuthorFemale1 -0.0834 0.0841 -0.99 0.3220
## LastAuthorFemale1 -0.0125 0.0818 -0.15 0.8787
## UniqueAuthors2 0.1440 0.0835 1.73 0.0855 .
## UniqueAuthors3 0.1551 0.1289 1.20 0.2298
## UniqueAuthors4 0.4419 0.1606 2.75 0.0063 **
## UniqueAuthors5 0.3354 0.1517 2.21 0.0277 *
## Year1997 -0.4714 0.2764 -1.71 0.0891 .
## Year1998 -0.1447 0.2669 -0.54 0.5881
## Year1999 -0.2663 0.2522 -1.06 0.2917
```

```

## Year2000          0.0233      0.3117      0.07      0.9405
## Year2001         -0.3317      0.3373     -0.98      0.3263
## Year2002         -0.0886      0.2603     -0.34      0.7338
## Year2003         -0.3459      0.2659     -1.30      0.1943
## Year2004         -0.3936      0.2962     -1.33      0.1848
## Year2005         -0.1627      0.2638     -0.62      0.5378
## Year2006         -0.0568      0.2657     -0.21      0.8308
## Year2007         -0.2471      0.2511     -0.98      0.3258
## Year2008         -0.1906      0.2385     -0.80      0.4248
## Year2009         -0.3264      0.2614     -1.25      0.2128
## Year2010         -0.3359      0.2487     -1.35      0.1778
## Year2011         -0.2418      0.2487     -0.97      0.3317
## Year2012         -0.4271      0.2611     -1.64      0.1029
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.621
## Multiple R-squared:  0.0724, Adjusted R-squared:  0.00698
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 313 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.443  0.880  0.955  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      2.99e-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 1.872 1      1.368
## LastAuthorFemale  1.764 1      1.328
## Year              1.451 16      1.012

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.2967 -0.3875 -0.0118 0.4319 1.6378
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5327 0.2363 6.49 3.4e-10 ***
## FirstAuthorFemale1 -0.0734 0.0885 -0.83 0.407
## LastAuthorFemale1 -0.0110 0.0857 -0.13 0.898
## Year1997 -0.5295 0.2874 -1.84 0.066 .
## Year1998 -0.2360 0.2747 -0.86 0.391
## Year1999 -0.3466 0.2643 -1.31 0.191
## Year2000 -0.0374 0.3267 -0.11 0.909
## Year2001 -0.4107 0.3603 -1.14 0.255
## Year2002 -0.1170 0.2692 -0.43 0.664
## Year2003 -0.4304 0.2775 -1.55 0.122
## Year2004 -0.4468 0.3010 -1.48 0.139
## Year2005 -0.1980 0.2704 -0.73 0.465
```

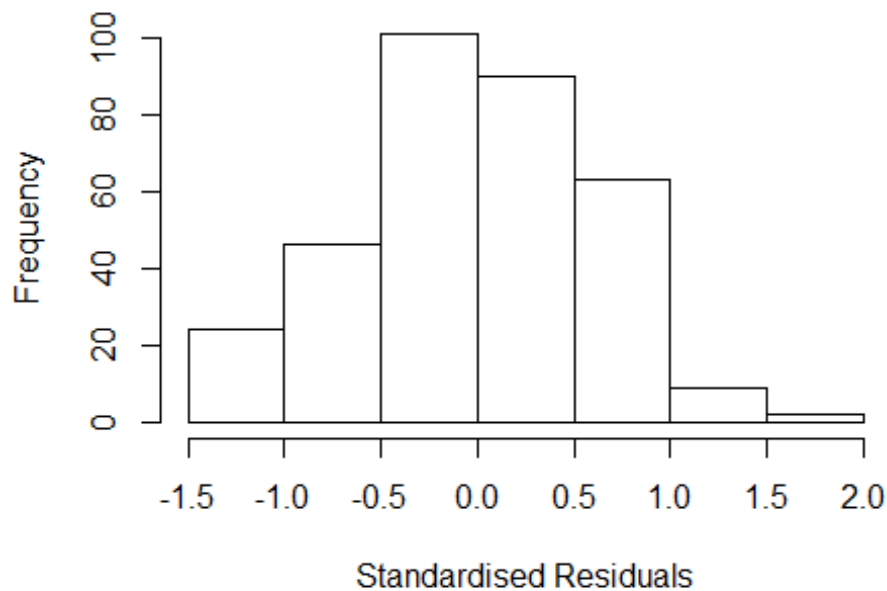


```

## Year2006          -0.1030      0.2740   -0.38    0.707
## Year2007          -0.3000      0.2586   -1.16    0.247
## Year2008          -0.2106      0.2560   -0.82    0.411
## Year2009          -0.3793      0.2693   -1.41    0.160
## Year2010          -0.3949      0.2655   -1.49    0.138
## Year2011          -0.2850      0.2622   -1.09    0.278
## Year2012          -0.4505      0.2819   -1.60    0.111
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.614
## Multiple R-squared:  0.0525, Adjusted R-squared:  -0.00152
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 305 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.457  0.870  0.950  0.913  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      2.99e-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 1.305 1          1.142
## Year              1.305 16          1.008

```

Residuals from first author



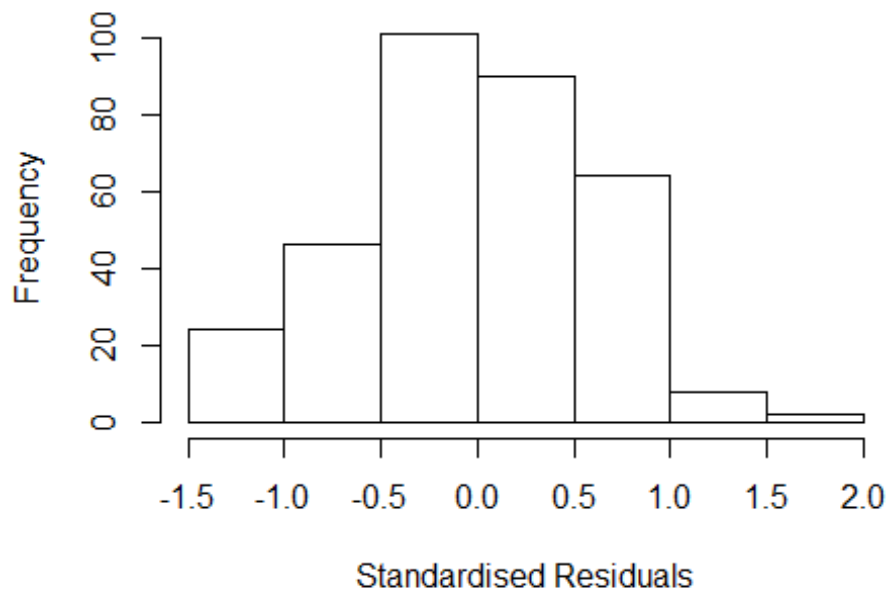
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2957 -0.3930 -0.0146 0.4312 1.6394
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5301 0.2371 6.45 4.1e-10 ***
## FirstAuthorFemale1 -0.0803 0.0739 -1.09 0.278
## Year1997 -0.5285 0.2867 -1.84 0.066 .
## Year1998 -0.2345 0.2757 -0.85 0.396
## Year1999 -0.3439 0.2647 -1.30 0.195
## Year2000 -0.0376 0.3250 -0.12 0.908
## Year2001 -0.4097 0.3595 -1.14 0.255
## Year2002 -0.1167 0.2688 -0.43 0.665
## Year2003 -0.4296 0.2773 -1.55 0.122
## Year2004 -0.4452 0.3014 -1.48 0.141
## Year2005 -0.1973 0.2700 -0.73 0.466
## Year2006 -0.1029 0.2734 -0.38 0.707
```

```

## Year2007          -0.2993      0.2585   -1.16    0.248
## Year2008          -0.2097      0.2558   -0.82    0.413
## Year2009          -0.3790      0.2684   -1.41    0.159
## Year2010          -0.3931      0.2656   -1.48    0.140
## Year2011          -0.2835      0.2632   -1.08    0.282
## Year2012          -0.4492      0.2822   -1.59    0.112
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.618
## Multiple R-squared:  0.0523, Adjusted R-squared:  0.00144
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 306 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.462  0.871  0.951  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      2.99e-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 1.232 1      1.110
## Year              1.232 16      1.007

```

Residuals from last author



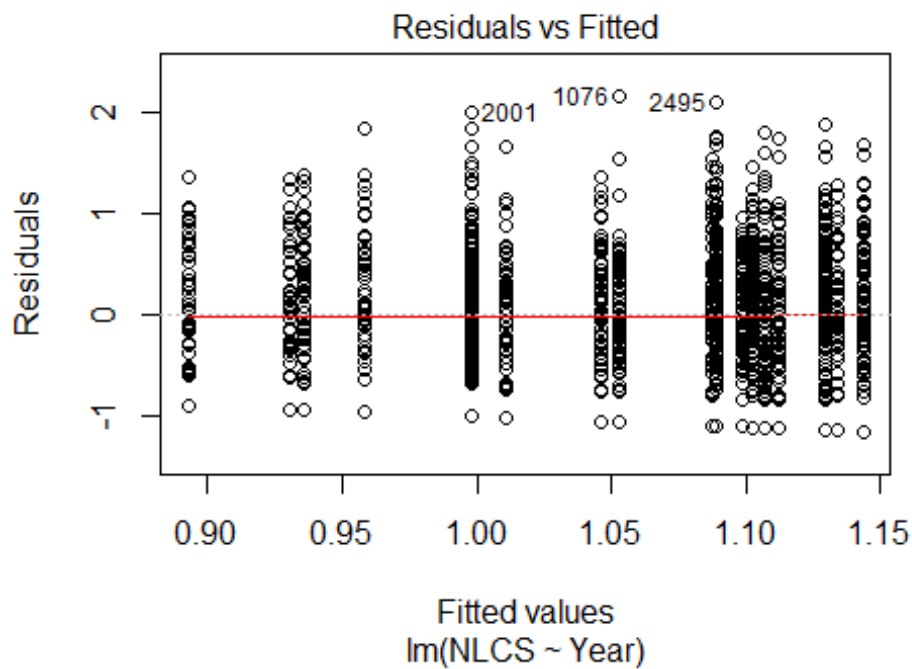
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2802 -0.3995 -0.0149 0.4260 1.6434
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5332 0.2361 6.49 3.2e-10 ***
## LastAuthorFemale1 -0.0574 0.0721 -0.80 0.426
## Year1997 -0.5357 0.2882 -1.86 0.064 .
## Year1998 -0.2530 0.2749 -0.92 0.358
## Year1999 -0.3569 0.2646 -1.35 0.178
## Year2000 -0.0472 0.3295 -0.14 0.886
## Year2001 -0.4301 0.3540 -1.21 0.225
## Year2002 -0.1256 0.2681 -0.47 0.640
## Year2003 -0.4387 0.2774 -1.58 0.115
## Year2004 -0.4679 0.2982 -1.57 0.118
## Year2005 -0.2046 0.2723 -0.75 0.453
## Year2006 -0.1200 0.2748 -0.44 0.663
```

```

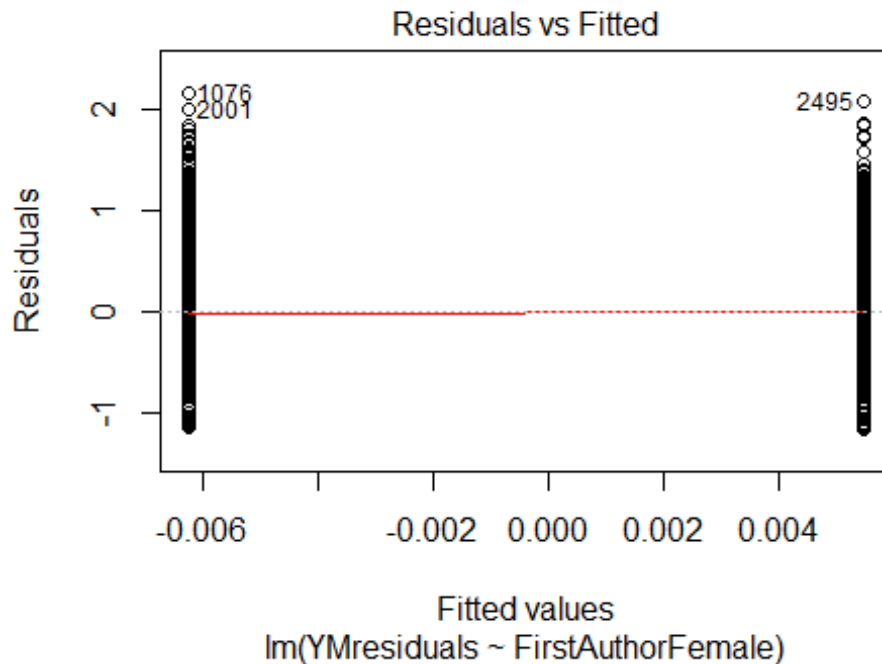
## Year2007          -0.3159      0.2583    -1.22      0.222
## Year2008          -0.2234      0.2567    -0.87      0.385
## Year2009          -0.3782      0.2700    -1.40      0.162
## Year2010          -0.4045      0.2655    -1.52      0.129
## Year2011          -0.3063      0.2625    -1.17      0.244
## Year2012          -0.4687      0.2808    -1.67      0.096 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.621
## Multiple R-squared:  0.0503, Adjusted R-squared:  -0.000626
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 26 weights are ~= 1. The remaining 309 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.463  0.879  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      2.99e-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: 335"
## [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
##   62   64   64   85   84   90   70   93   87  108  110  129  142  196  248
## 2011 2012
##  345  273
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   56   47   54   73   72   70   62   82   73   99   89   96  119  159  216
## 2011 2012

```

```
## 283 231
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 55 46 54 71 72 64 59 77 70 95 88 89 110 151 202
## 2011 2012
## 262 216
## [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
```

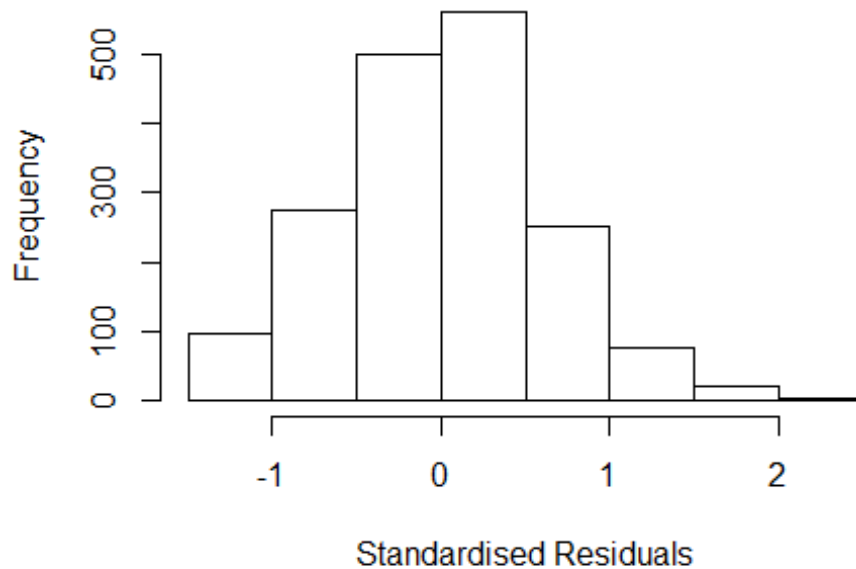


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 9, df = 1, p-value = 0.003
```



```
## [1] "Female first author team size 2018 geometric mean: 1.87136490225644"
## [1] "Male first author team size 2018 geometric mean: 1.64593887324879"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5500, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.82727637740648"
## [1] "Male last author team size 2018 geometric mean: 1.74194194777957"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 5500, 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 1.762 1          1.327
## LastAuthorFemale  1.755 1          1.325
## UniqueAuthors    1.179 4          1.021
## Year              1.254 16         1.007
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.3373 -0.3995 0.0179 0.3904 2.2625
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.84976 0.09812 8.66 < 2e-16 ***
## FirstAuthorFemale1 -0.00464 0.03929 -0.12 0.9059
## LastAuthorFemale1 0.03530 0.03892 0.91 0.3646
## UniqueAuthors2 0.09419 0.03596 2.62 0.0089 **
## UniqueAuthors3 0.20538 0.04933 4.16 3.3e-05 ***
## UniqueAuthors4 0.18393 0.07074 2.60 0.0094 **
## UniqueAuthors5 0.32482 0.08190 3.97 7.6e-05 ***
## Year1997 0.21256 0.14103 1.51 0.1319
## Year1998 0.01995 0.12893 0.15 0.8770
## Year1999 0.06926 0.11829 0.59 0.5583
```

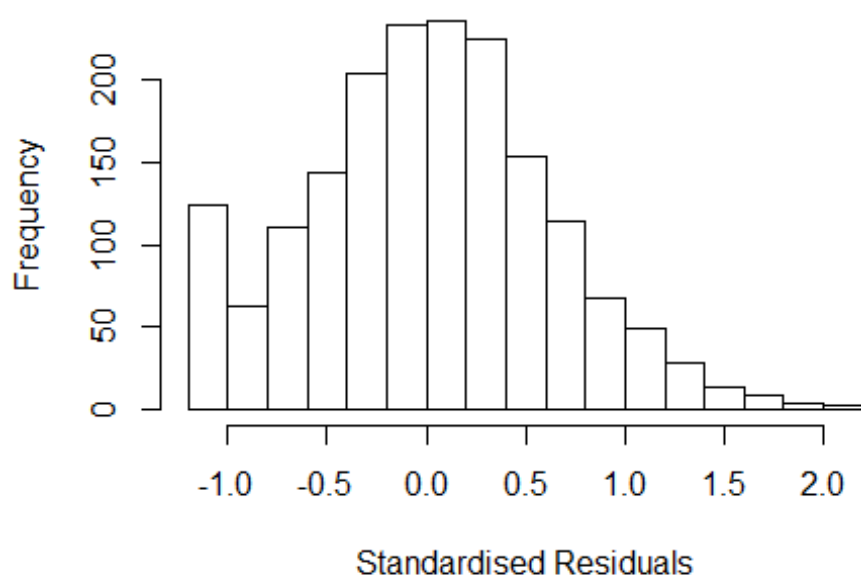


```

## Year2000      0.00328    0.12621    0.03    0.9793
## Year2001      0.12638    0.12966    0.97    0.3298
## Year2002     -0.00347    0.13193   -0.03    0.9790
## Year2003      0.08946    0.12560    0.71    0.4764
## Year2004      0.14435    0.12348    1.17    0.2426
## Year2005      0.17931    0.11538    1.55    0.1203
## Year2006      0.08775    0.11364    0.77    0.4401
## Year2007      0.19380    0.11084    1.75    0.0806 .
## Year2008      0.18612    0.11493    1.62    0.1055
## Year2009      0.18223    0.10769    1.69    0.0908 .
## Year2010      0.21114    0.10738    1.97    0.0494 *
## Year2011      0.06333    0.10416    0.61    0.5433
## Year2012      0.13211    0.10729    1.23    0.2183
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.599
## Multiple R-squared:  0.0328, Adjusted R-squared:  0.0207
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 157 weights are ~= 1. The remaining 1624 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.123  0.863  0.951  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      5.61e-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 1.791 1      1.338
## LastAuthorFemale  1.801 1      1.342
## Year              1.101 16      1.003

```

Residuals from first and last author



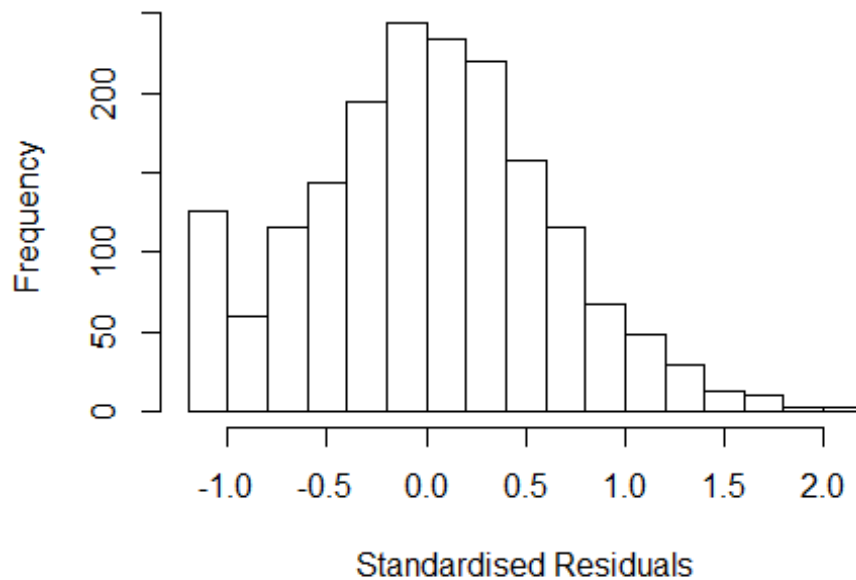
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.15276 -0.39765 0.00897 0.39592 2.17859
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.87159 0.09860 8.84 <2e-16 ***
## FirstAuthorFemale1 0.00302 0.04015 0.08 0.940
## LastAuthorFemale1 0.02864 0.03996 0.72 0.474
## Year1997 0.23539 0.14111 1.67 0.095 .
## Year1998 0.02788 0.12898 0.22 0.829
## Year1999 0.08824 0.11896 0.74 0.458
## Year2000 0.04041 0.12669 0.32 0.750
## Year2001 0.15584 0.13337 1.17 0.243
## Year2002 0.03424 0.13459 0.25 0.799
## Year2003 0.12643 0.12631 1.00 0.317
## Year2004 0.18172 0.12464 1.46 0.145
## Year2005 0.20269 0.11606 1.75 0.081 .
```

```

## Year2006          0.14982    0.11371    1.32    0.188
## Year2007          0.22138    0.11150    1.99    0.047 *
## Year2008          0.23382    0.11543    2.03    0.043 *
## Year2009          0.22458    0.10810    2.08    0.038 *
## Year2010          0.24950    0.10751    2.32    0.020 *
## Year2011          0.10148    0.10461    0.97    0.332
## Year2012          0.18203    0.10742    1.69    0.090 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.015, Adjusted R-squared:  0.00495
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 155 weights are ~= 1. The remaining 1626 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.167  0.864  0.951  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      5.61e-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 1.056 1      1.028
## Year              1.056 16      1.002

```

Residuals from first author



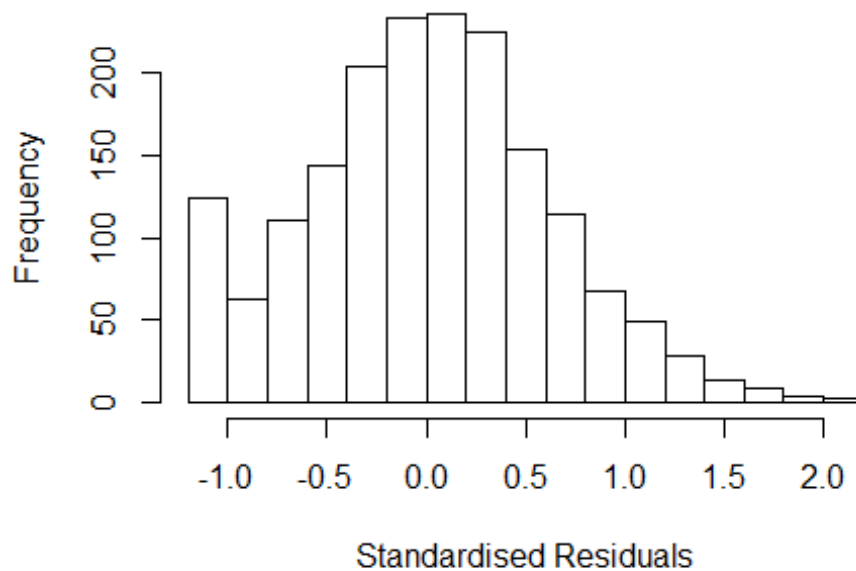
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.15027 -0.39660  0.00809  0.39702  2.17238
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8739    0.0987    8.86  <2e-16 ***
## FirstAuthorFemale1 0.0221    0.0309    0.72   0.475
## Year1997        0.2364    0.1412    1.67   0.094 .
## Year1998        0.0283    0.1291    0.22   0.826
## Year1999        0.0887    0.1192    0.74   0.457
## Year2000        0.0419    0.1271    0.33   0.742
## Year2001        0.1565    0.1333    1.17   0.241
## Year2002        0.0347    0.1343    0.26   0.796
## Year2003        0.1289    0.1263    1.02   0.308
## Year2004        0.1847    0.1245    1.48   0.138
## Year2005        0.2033    0.1163    1.75   0.081 .
## Year2006        0.1538    0.1138    1.35   0.177
```

```

## Year2007          0.2261      0.1116      2.03      0.043 *
## Year2008          0.2365      0.1155      2.05      0.041 *
## Year2009          0.2250      0.1083      2.08      0.038 *
## Year2010          0.2543      0.1075      2.37      0.018 *
## Year2011          0.1050      0.1048      1.00      0.317
## Year2012          0.1842      0.1077      1.71      0.087 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.0148, Adjusted R-squared:  0.00525
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 157 weights are ~= 1. The remaining 1624 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.169  0.864  0.951   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      5.61e-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 1.063 1      1.031
## Year            1.063 16      1.002

```

Residuals from last author



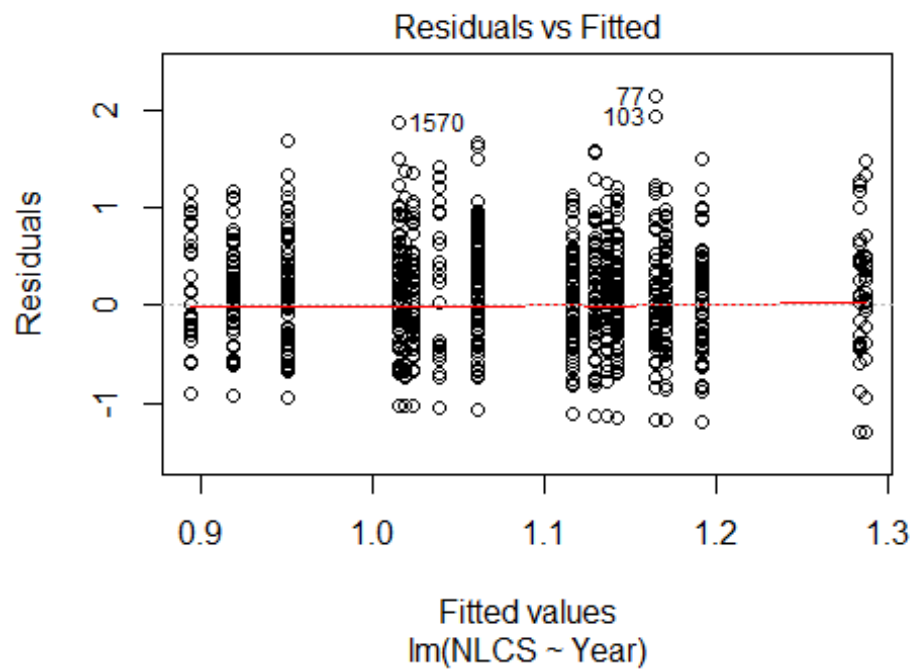
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.15226 -0.39782 0.00843 0.39652 2.17818
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8721 0.0983 8.87 <2e-16 ***
## LastAuthorFemale1 0.0307 0.0307 1.00 0.318
## Year1997 0.2352 0.1411 1.67 0.096 .
## Year1998 0.0277 0.1290 0.22 0.830
## Year1999 0.0883 0.1190 0.74 0.458
## Year2000 0.0404 0.1267 0.32 0.750
## Year2001 0.1560 0.1334 1.17 0.242
## Year2002 0.0341 0.1346 0.25 0.800
## Year2003 0.1265 0.1263 1.00 0.317
## Year2004 0.1817 0.1247 1.46 0.145
## Year2005 0.2029 0.1159 1.75 0.080 .
## Year2006 0.1497 0.1137 1.32 0.188
```

```

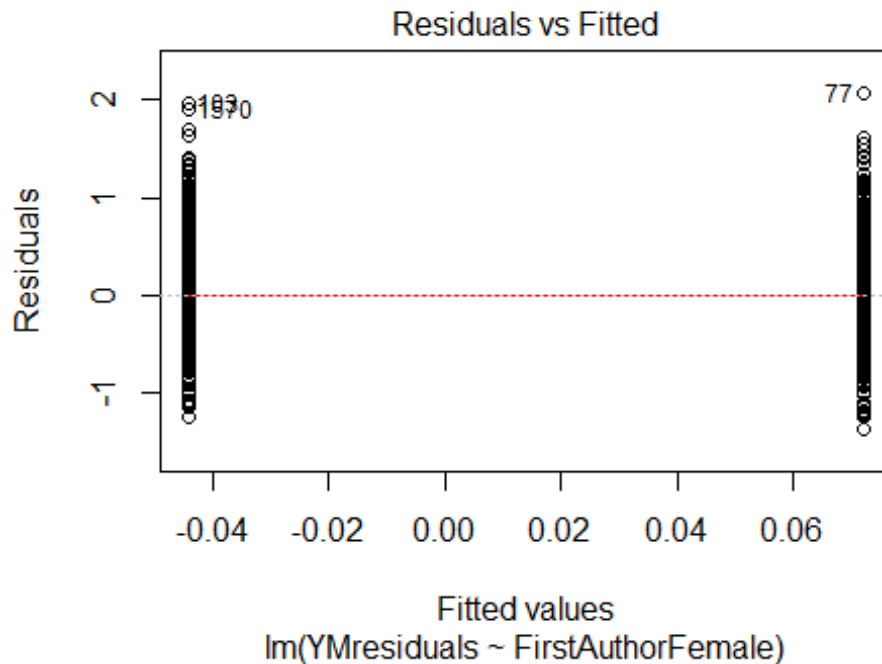
## Year2007          0.2213      0.1115      1.99      0.047 *
## Year2008          0.2339      0.1154      2.03      0.043 *
## Year2009          0.2249      0.1081      2.08      0.038 *
## Year2010          0.2495      0.1075      2.32      0.020 *
## Year2011          0.1016      0.1046      0.97      0.331
## Year2012          0.1822      0.1073      1.70      0.090 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.015, Adjusted R-squared:  0.00551
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 155 weights are ~= 1. The remaining 1626 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.166  0.863  0.951   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      5.61e-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: 1781"
## [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
##   42   64   51   74   65   82   78  104   77   81   92   86  100   97  104
## 2011 2012
##  140  125
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   30   49   33   45   31   30   58   88   62   72   75   77   83   70   84
## 2011 2012

```

```
## 110 106
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 27 48 33 41 29 28 55 85 58 69 70 73 78 62 82
## 2011 2012
## 104 100
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 36, df = 16, p-value = 0.003
```

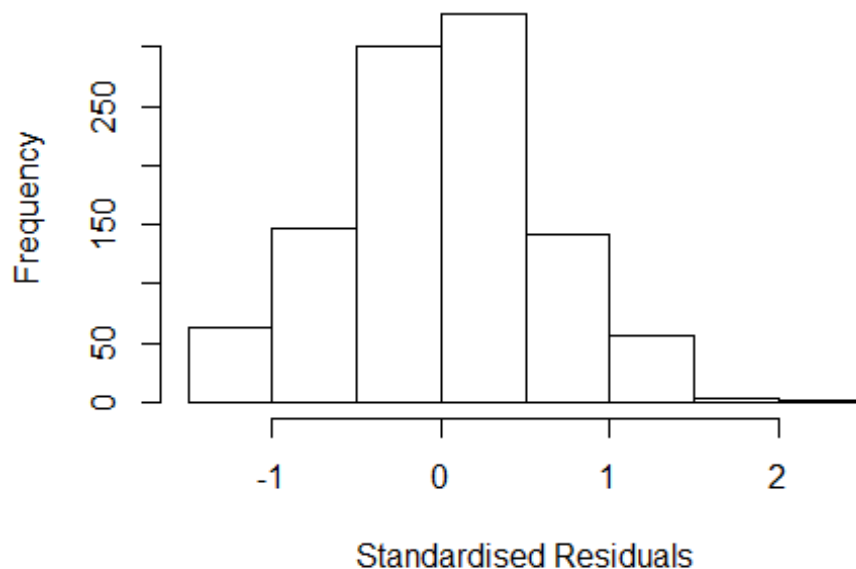


```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.21, df = 1, p-value = 0.6
```

```
## [1] "Female first author team size 2018 geometric mean: 1.73077235750838"
## [1] "Male first author team size 2018 geometric mean: 1.75438759074303"
##
## 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: 1.84173509019665"
## [1] "Male last author team size 2018 geometric mean: 1.64017202455831"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1400, 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 1.751 1          1.323
## LastAuthorFemale  1.763 1          1.328
## UniqueAuthors    1.311 4          1.034
## Year              1.477 16         1.012
```

Residuals from first and last author and team size



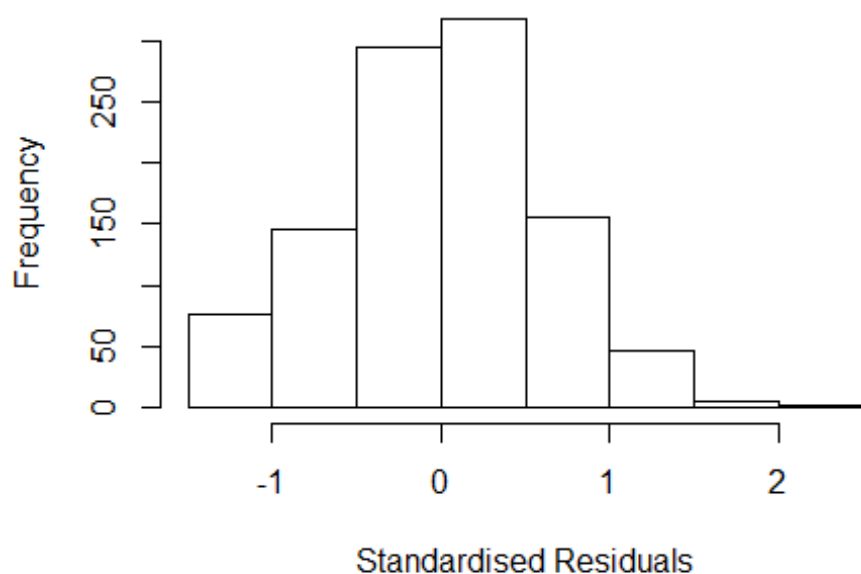
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3983 -0.4101 0.0149 0.3903 2.1469
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2706 0.1493 8.51 <2e-16 ***
## FirstAuthorFemale1 0.1432 0.0520 2.76 0.0059 **
## LastAuthorFemale1 -0.0156 0.0528 -0.29 0.7684
## UniqueAuthors2 0.1576 0.0488 3.23 0.0013 **
## UniqueAuthors3 0.1532 0.0623 2.46 0.0141 *
## UniqueAuthors4 0.2382 0.0923 2.58 0.0100 *
## UniqueAuthors5 0.1822 0.1281 1.42 0.1554
## Year1997 -0.2432 0.1855 -1.31 0.1900
## Year1998 -0.4575 0.2032 -2.25 0.0246 *
## Year1999 -0.3255 0.1769 -1.84 0.0661 .
```

```

## Year2000          -0.3193      0.2435   -1.31   0.1900
## Year2001          -0.0351      0.2082   -0.17   0.8661
## Year2002          -0.2306      0.1764   -1.31   0.1913
## Year2003          -0.4123      0.1699   -2.43   0.0154 *
## Year2004          -0.4570      0.1706   -2.68   0.0075 **
## Year2005          -0.2044      0.1671   -1.22   0.2216
## Year2006          -0.2814      0.1641   -1.72   0.0866 .
## Year2007          -0.2029      0.1663   -1.22   0.2228
## Year2008          -0.3019      0.1655   -1.82   0.0685 .
## Year2009          -0.3237      0.1642   -1.97   0.0489 *
## Year2010          -0.2608      0.1595   -1.64   0.1023
## Year2011          -0.3857      0.1636   -2.36   0.0186 *
## Year2012          -0.3287      0.1678   -1.96   0.0504 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.596
## Multiple R-squared:  0.0538, Adjusted R-squared:  0.0333
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 99 weights are ~= 1. The remaining 943 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.167  0.852  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          9.60e-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 1.710 1          1.308
## LastAuthorFemale  1.736 1          1.318
## Year              1.186 16          1.005

```

Residuals from first and last author



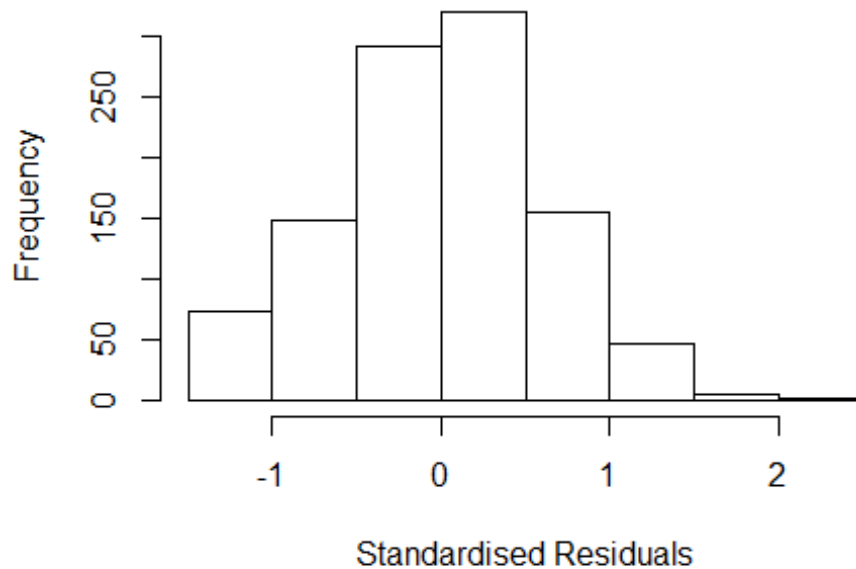
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4663 -0.3917 0.0147 0.4121 2.0846
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3388 0.1566 8.55 <2e-16 ***
## FirstAuthorFemale1 0.1546 0.0524 2.95 0.0032 **
## LastAuthorFemale1 -0.0270 0.0534 -0.51 0.6134
## Year1997 -0.2490 0.1915 -1.30 0.1939
## Year1998 -0.4992 0.2102 -2.37 0.0178 *
## Year1999 -0.3376 0.1828 -1.85 0.0650 .
## Year2000 -0.3743 0.2522 -1.48 0.1380
## Year2001 -0.0521 0.2175 -0.24 0.8108
## Year2002 -0.2291 0.1846 -1.24 0.2149
## Year2003 -0.4361 0.1777 -2.45 0.0143 *
## Year2004 -0.4821 0.1776 -2.71 0.0067 **
## Year2005 -0.2148 0.1742 -1.23 0.2177
```

```

## Year2006          -0.2929      0.1719    -1.70    0.0887 .
## Year2007          -0.2130      0.1742    -1.22    0.2218
## Year2008          -0.2968      0.1738    -1.71    0.0880 .
## Year2009          -0.3396      0.1732    -1.96    0.0502 .
## Year2010          -0.2624      0.1677    -1.56    0.1179
## Year2011          -0.3808      0.1711    -2.23    0.0262 *
## Year2012          -0.3247      0.1754    -1.85    0.0644 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.603
## Multiple R-squared:  0.0384, Adjusted R-squared:  0.0214
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 84 weights are ~= 1. The remaining 958 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.207  0.853   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      9.60e-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 1.114 1          1.056
## Year              1.114 16          1.003

```

Residuals from first author



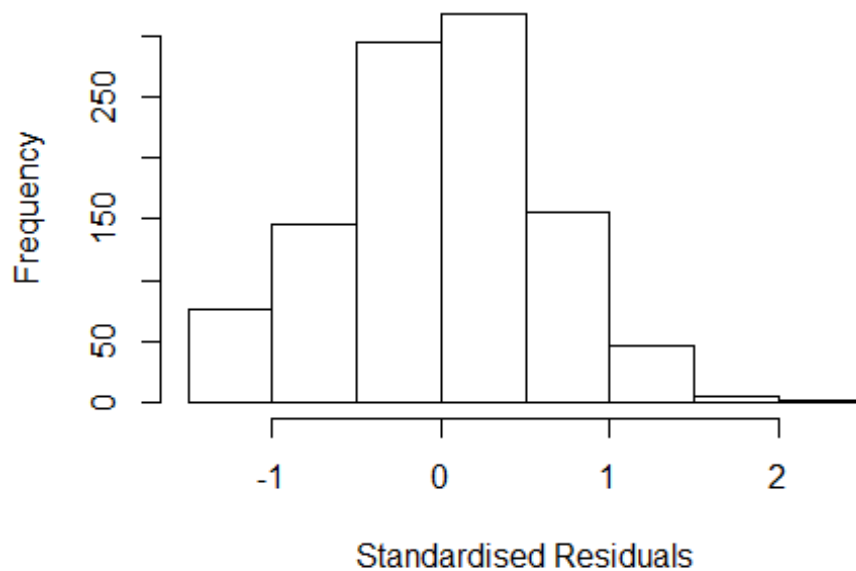
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4728 -0.3919 0.0128 0.4112 2.0789
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3363 0.1571 8.51 <2e-16 ***
## FirstAuthorFemale1 0.1365 0.0425 3.21 0.0013 **
## Year1997 -0.2497 0.1913 -1.31 0.1922
## Year1998 -0.4982 0.2104 -2.37 0.0181 *
## Year1999 -0.3367 0.1830 -1.84 0.0660 .
## Year2000 -0.3728 0.2503 -1.49 0.1368
## Year2001 -0.0564 0.2162 -0.26 0.7941
## Year2002 -0.2304 0.1841 -1.25 0.2110
## Year2003 -0.4378 0.1772 -2.47 0.0136 *
## Year2004 -0.4835 0.1772 -2.73 0.0065 **
## Year2005 -0.2151 0.1741 -1.24 0.2170
## Year2006 -0.2927 0.1719 -1.70 0.0889 .
```

```

## Year2007          -0.2139      0.1739   -1.23   0.2192
## Year2008          -0.2949      0.1740   -1.69   0.0904 .
## Year2009          -0.3401      0.1731   -1.97   0.0497 *
## Year2010          -0.2617      0.1679   -1.56   0.1193
## Year2011          -0.3810      0.1710   -2.23   0.0261 *
## Year2012          -0.3254      0.1752   -1.86   0.0635 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.604
## Multiple R-squared:  0.0382, Adjusted R-squared:  0.0222
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 88 weights are ~= 1. The remaining 954 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.212  0.852  0.951  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      9.60e-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 1.132 1          1.064
## Year            1.132 16          1.004

```

Residuals from last author



```
## [1] "List of 0 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
## -1.4209 -0.4126 0.0132 0.4098 2.1351
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3423 0.1564 8.58 <2e-16 ***
## LastAuthorFemale1 0.0785 0.0433 1.82 0.0698 .
## Year1997 -0.2540 0.1922 -1.32 0.1866
## Year1998 -0.4847 0.2110 -2.30 0.0218 *
## Year1999 -0.3241 0.1830 -1.77 0.0769 .
## Year2000 -0.3601 0.2575 -1.40 0.1623
## Year2001 -0.0497 0.2181 -0.23 0.8197
## Year2002 -0.2132 0.1834 -1.16 0.2453
## Year2003 -0.4215 0.1776 -2.37 0.0178 *
## Year2004 -0.4679 0.1767 -2.65 0.0082 **
## Year2005 -0.1902 0.1738 -1.09 0.2740
## Year2006 -0.2706 0.1718 -1.58 0.1155
```

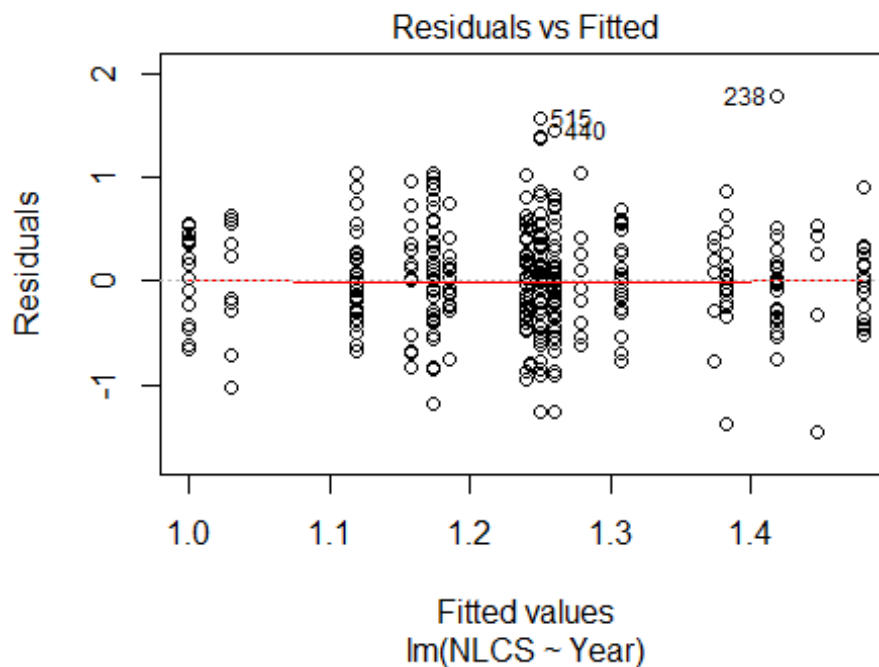


```

## Year2007          -0.1940      0.1740   -1.12   0.2650
## Year2008          -0.2667      0.1729   -1.54   0.1234
## Year2009          -0.3261      0.1730   -1.88   0.0598 .
## Year2010          -0.2442      0.1672   -1.46   0.1444
## Year2011          -0.3561      0.1708   -2.09   0.0373 *
## Year2012          -0.3120      0.1759   -1.77   0.0765 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.608
## Multiple R-squared:  0.0305, Adjusted R-squared:  0.0144
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 87 weights are ~= 1. The remaining 955 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.192  0.860  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      9.60e-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: 1042"
## [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
##    9   20   10   19   18   25   14   18   19   22   24   31   38   41   51
## 2011 2012
##   51   57
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6   17    6   15   10   16    9   16   14   16   21   25   30   32   43
## 2011 2012

```

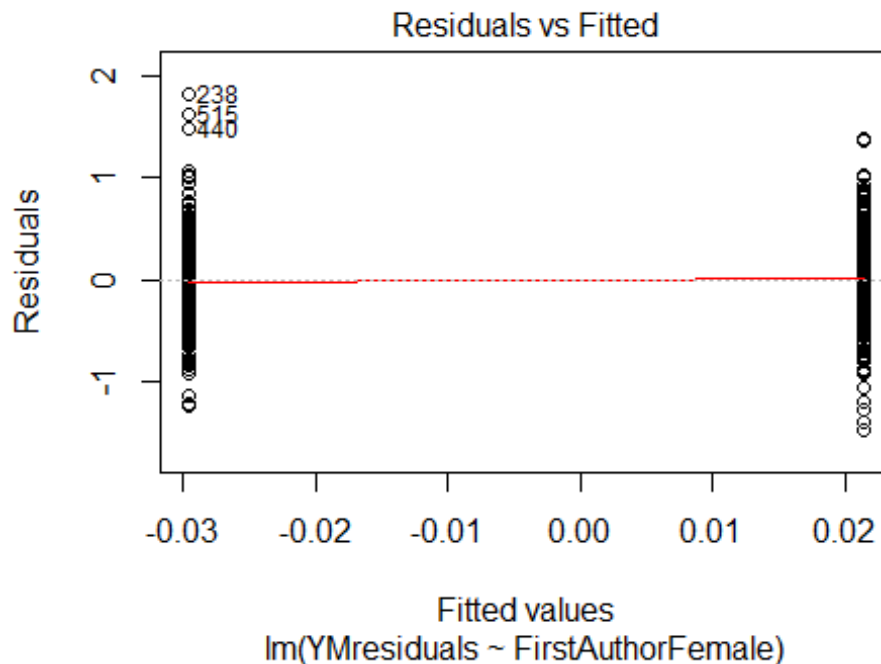
```
## 47 47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 17 5 14 10 15 8 12 13 15 19 22 28 27 40
## 2011 2012
## 43 42
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 0.87, df = 1, p-value = 0.4
## [1] "Female first author team size 2018 geometric mean: 2.70432249553603"
## [1] "Male first author team size 2018 geometric mean: 2.17733143766132"
## 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 = 250, p-value = 0.4
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.5388679963372"
## [1] "Male last author team size 2018 geometric mean: 2.40457535699717"

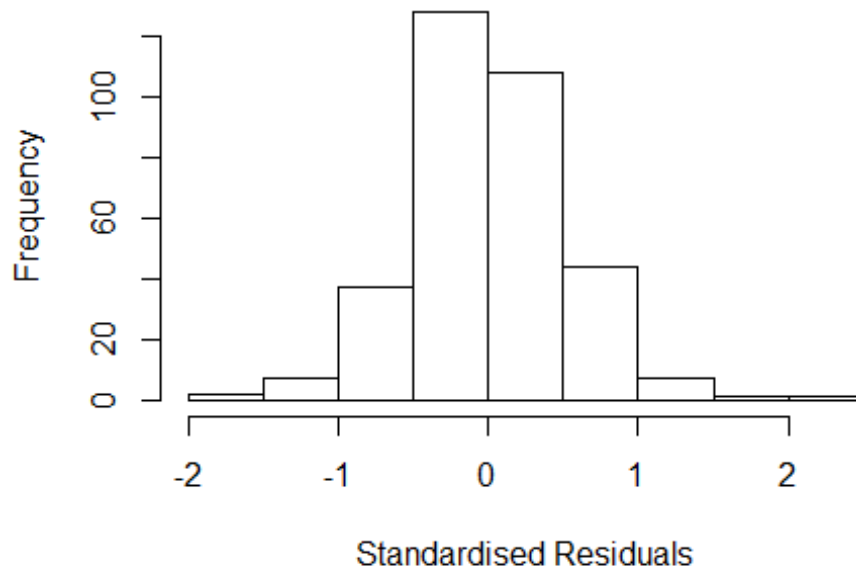
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 230, 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	1.735	1	1.317
LastAuthorFemale	1.966	1	1.402
UniqueAuthors	2.961	4	1.145
Year	4.321	16	1.047

Residuals from first and last author and team size



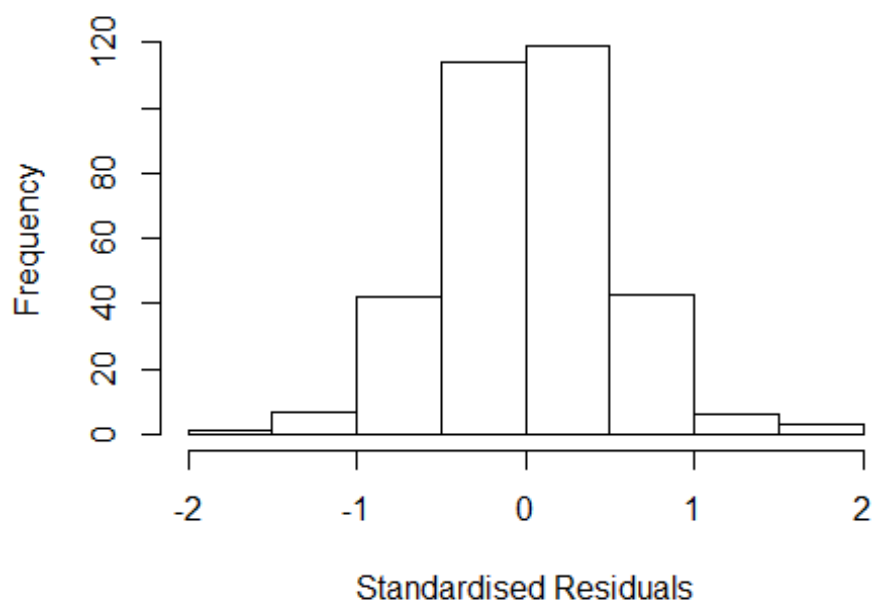
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.686 -0.343 -0.033 0.333 2.040
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4725 0.3070 4.80 2.5e-06 ***
## FirstAuthorFemale1 0.1047 0.0699 1.50 0.13531
## LastAuthorFemale1 -0.0480 0.0729 -0.66 0.51136
## UniqueAuthors2 0.0996 0.0804 1.24 0.21674
## UniqueAuthors3 0.2614 0.0766 3.41 0.00073 ***
## UniqueAuthors4 0.3047 0.1185 2.57 0.01063 *
## UniqueAuthors5 0.2253 0.1062 2.12 0.03471 *
## Year1997 -0.6249 0.3276 -1.91 0.05734 .
## Year1998 -0.2438 0.3705 -0.66 0.51098
## Year1999 -0.1476 0.3233 -0.46 0.64832
```

```

## Year2000          -0.6419      0.3538    -1.81   0.07063 .
## Year2001          -0.1178      0.3270    -0.36   0.71890
## Year2002          -0.3565      0.3751    -0.95   0.34265
## Year2003          -0.3907      0.3420    -1.14   0.25426
## Year2004          -0.3847      0.3364    -1.14   0.25367
## Year2005          -0.4971      0.3398    -1.46   0.14449
## Year2006          -0.3129      0.3178    -0.98   0.32552
## Year2007          -0.3187      0.3295    -0.97   0.33422
## Year2008          -0.5403      0.3269    -1.65   0.09935 .
## Year2009          -0.4271      0.3231    -1.32   0.18709
## Year2010          -0.4339      0.3281    -1.32   0.18691
## Year2011          -0.3872      0.3254    -1.19   0.23495
## Year2012          -0.3990      0.3294    -1.21   0.22679
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.517
## Multiple R-squared:  0.102, Adjusted R-squared:  0.0386
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 311 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0847 0.8940 0.9560 0.9130 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          2.99e-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 1.585 1          1.259
## LastAuthorFemale 1.711 1          1.308
## Year          1.759 16          1.018

```

Residuals from first and last author



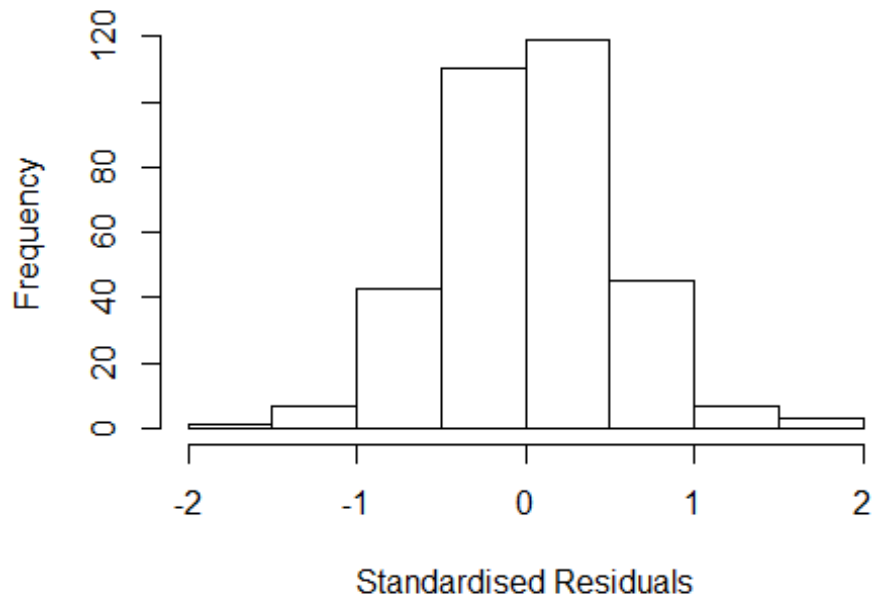
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.5992 -0.3367  0.0131  0.3269  1.8804
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.5960     0.3273   4.88 1.7e-06 ***
## FirstAuthorFemale1  0.1139     0.0683   1.67  0.096 .
## LastAuthorFemale1 -0.1106     0.0702  -1.57  0.116
## Year1997          -0.6144     0.3508  -1.75  0.081 .
## Year1998          -0.2459     0.3873  -0.64  0.526
## Year1999          -0.1856     0.3512  -0.53  0.597
## Year2000          -0.5738     0.3819  -1.50  0.134
## Year2001          -0.1430     0.3494  -0.41  0.683
## Year2002          -0.3491     0.3881  -0.90  0.369
## Year2003          -0.3674     0.3631  -1.01  0.312
## Year2004          -0.3993     0.3558  -1.12  0.263
## Year2005          -0.4444     0.3740  -1.19  0.236
```

```

## Year2006          -0.2764      0.3446   -0.80    0.423
## Year2007          -0.3080      0.3514   -0.88    0.381
## Year2008          -0.4518      0.3495   -1.29    0.197
## Year2009          -0.3636      0.3491   -1.04    0.298
## Year2010          -0.4232      0.3524   -1.20    0.231
## Year2011          -0.3976      0.3498   -1.14    0.257
## Year2012          -0.3853      0.3533   -1.09    0.276
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.0571, Adjusted R-squared:  0.00343
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 303 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.162  0.883  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      2.99e-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 1.328 1          1.152
## Year              1.328 16          1.009

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.6401 -0.3330 0.0255 0.3360 1.9207
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5724 0.3204 4.91 1.5e-06 ***
## FirstAuthorFemale1 0.0676 0.0626 1.08 0.281
## Year1997 -0.6107 0.3441 -1.77 0.077 .
## Year1998 -0.2654 0.3815 -0.70 0.487
## Year1999 -0.1757 0.3442 -0.51 0.610
## Year2000 -0.5473 0.3791 -1.44 0.150
## Year2001 -0.1408 0.3433 -0.41 0.682
## Year2002 -0.3684 0.3913 -0.94 0.347
## Year2003 -0.3640 0.3556 -1.02 0.307
## Year2004 -0.4301 0.3504 -1.23 0.221
## Year2005 -0.4766 0.3630 -1.31 0.190
## Year2006 -0.2931 0.3380 -0.87 0.386
```

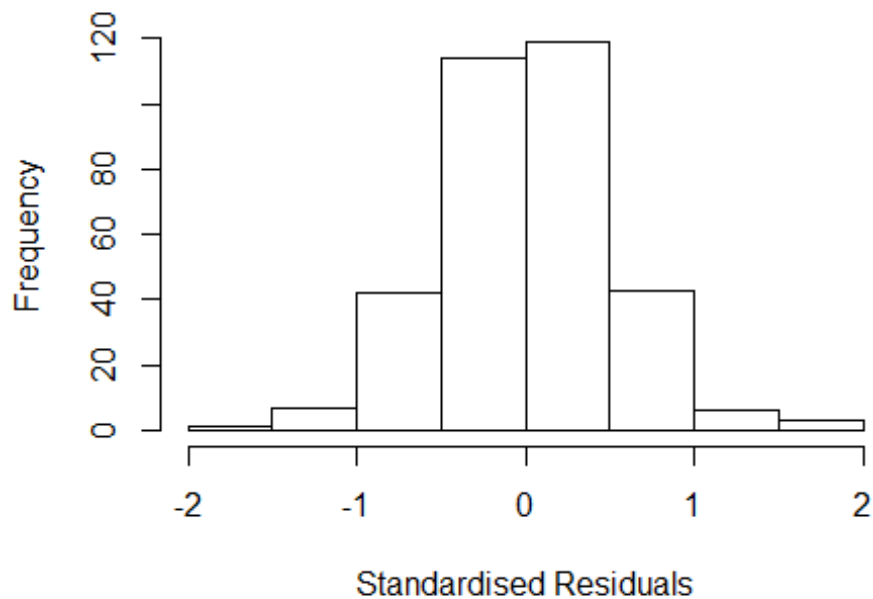


```

## Year2007          -0.3029      0.3438   -0.88    0.379
## Year2008          -0.4835      0.3419   -1.41    0.158
## Year2009          -0.3657      0.3420   -1.07    0.286
## Year2010          -0.4512      0.3447   -1.31    0.192
## Year2011          -0.3964      0.3451   -1.15    0.252
## Year2012          -0.3968      0.3442   -1.15    0.250
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.52
## Multiple R-squared:  0.0497, Adjusted R-squared:  -0.00124
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 29 weights are ~= 1. The remaining 306 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.142  0.888  0.954  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      2.99e-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 1.501 1          1.225
## Year            1.501 16          1.013

```

Residuals from last author



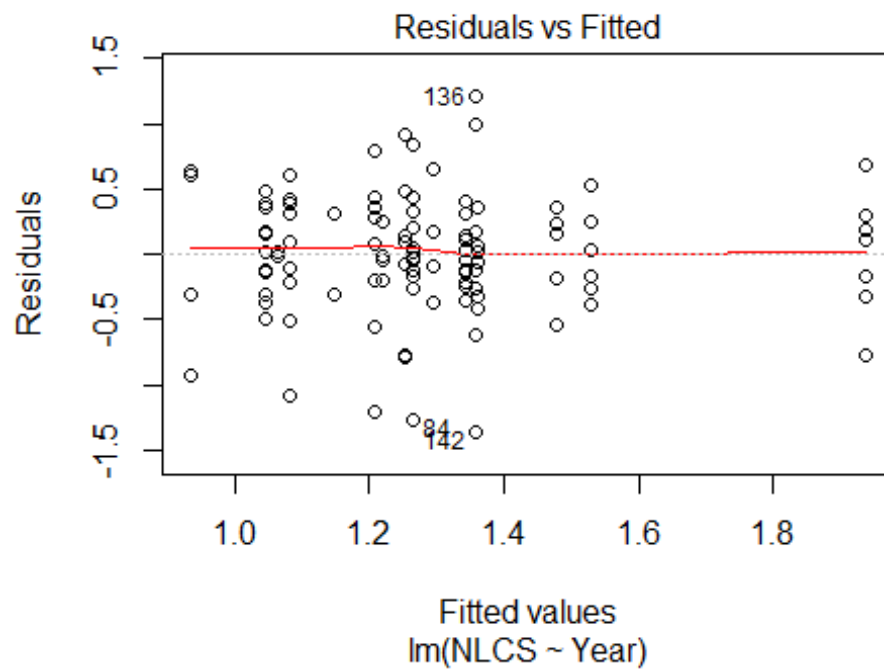
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.5030 -0.3392 0.0187 0.3395 1.8512
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5665 0.3278 4.78 2.7e-06 ***
## LastAuthorFemale1 -0.0635 0.0660 -0.96 0.34
## Year1997 -0.5355 0.3460 -1.55 0.12
## Year1998 -0.1730 0.3949 -0.44 0.66
## Year1999 -0.1302 0.3489 -0.37 0.71
## Year2000 -0.4908 0.3737 -1.31 0.19
## Year2001 -0.0712 0.3439 -0.21 0.84
## Year2002 -0.2701 0.3859 -0.70 0.48
## Year2003 -0.2951 0.3541 -0.83 0.41
## Year2004 -0.3056 0.3480 -0.88 0.38
## Year2005 -0.3962 0.3713 -1.07 0.29
## Year2006 -0.2177 0.3438 -0.63 0.53
```

```

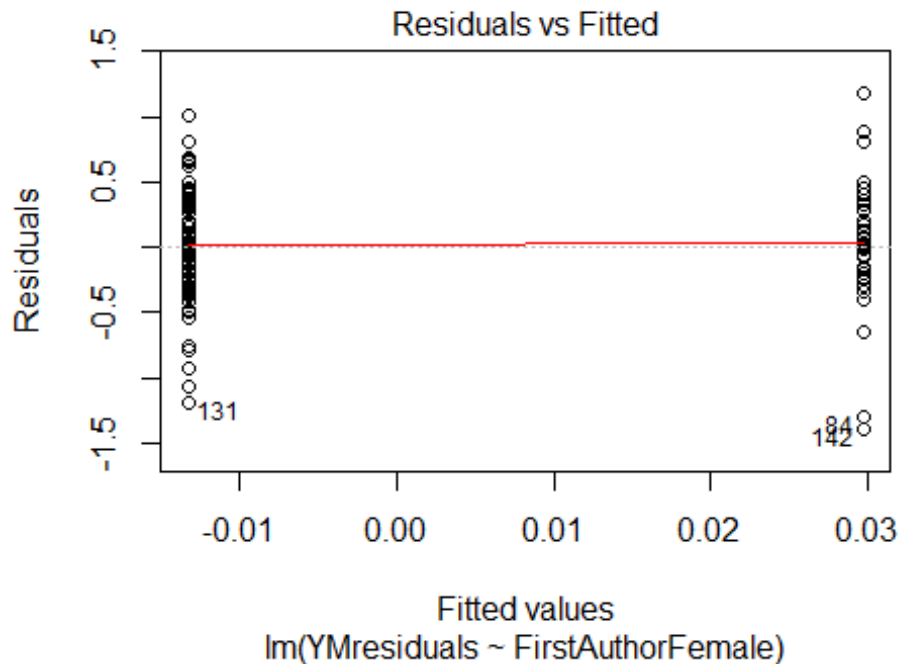
## Year2007          -0.2197      0.3448   -0.64      0.52
## Year2008          -0.3874      0.3454   -1.12      0.26
## Year2009          -0.3042      0.3439   -0.88      0.38
## Year2010          -0.3584      0.3497   -1.02      0.31
## Year2011          -0.3076      0.3435   -0.90      0.37
## Year2012          -0.3173      0.3483   -0.91      0.36
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.518
## Multiple R-squared:  0.0474, Adjusted R-squared:  -0.00366
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 310 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.175  0.884  0.956  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      2.99e-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: 335"
## [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
##    2    7    8    3    5    9    9    8   13    4    7    9   13    8   14
## 2011 2012
##    9   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    5    7    2    3    7    4    5   12    4    7    6    8    7   12
## 2011 2012

```

```
##      7    12
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      5      7      1      3      6      3      4     12      3      6      5      7      7      9
## 2011 2012
##      7      9
## [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
```



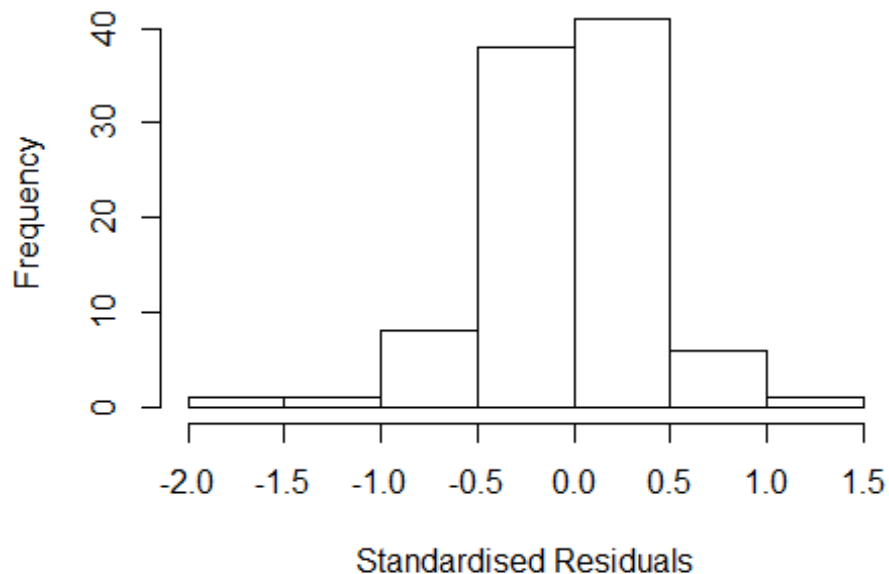
```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 1.5, df = 1, p-value = 0.2
```



```
## [1] "Female first author team size 2018 geometric mean: 6"
## [1] "Male first author team size 2018 geometric mean: 1.81712059283214"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 3, p-value = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3"
## [1] "Male last author team size 2018 geometric mean: 2.28942848510666"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 2, 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	14.439	1	3.800
## LastAuthorFemale	2.691	1	1.640
## UniqueAuthors	804.809	4	2.308
## Year	2713.958	16	1.280

Residuals from first and last author and team size



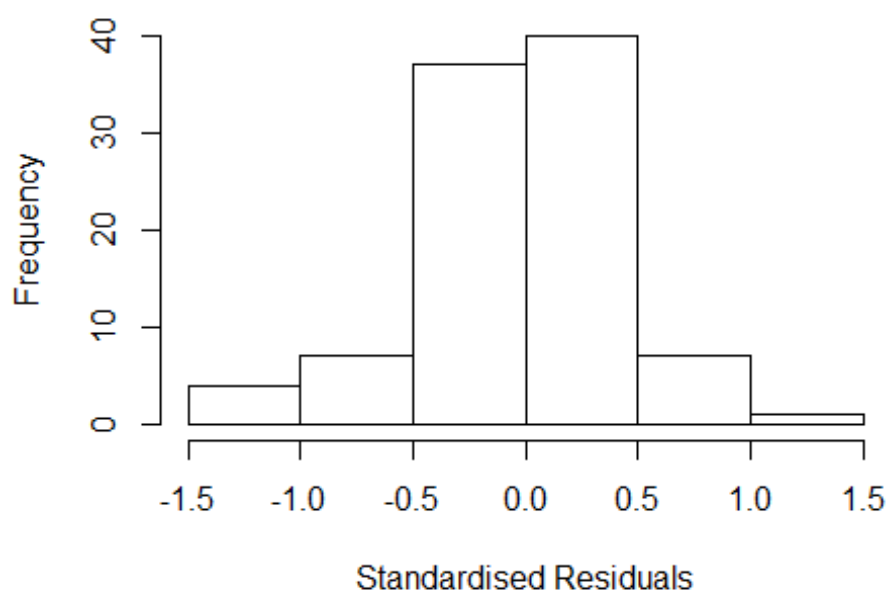
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.627197 -0.260759 0.000158 0.267722 1.137237
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9602 0.0975 9.85 4.8e-15 ***
## FirstAuthorFemale1 0.0369 0.1174 0.31 0.7541
## LastAuthorFemale1 -0.0582 0.1543 -0.38 0.7074
## UniqueAuthors2 0.2106 0.1520 1.39 0.1700
## UniqueAuthors3 0.4428 0.1841 2.40 0.0187 *
## UniqueAuthors4 0.5869 0.2310 2.54 0.0132 *
## UniqueAuthors5 0.2865 0.1755 1.63 0.1069
## Year1997 0.0366 0.2324 0.16 0.8754
## Year1998 0.1338 0.1975 0.68 0.5002
## Year1999 -0.1230 0.2728 -0.45 0.6533
```

```

## Year2000          -0.2013      0.1519    -1.32    0.1893
## Year2001           0.6849      0.2127     3.22    0.0019 **
## Year2002          -0.1583      0.4314    -0.37    0.7147
## Year2003           0.2503      0.1562     1.60    0.1132
## Year2004           0.1873      0.1852     1.01    0.3152
## Year2005          -0.2157      0.1672    -1.29    0.2010
## Year2006           0.1094      0.1700     0.64    0.5217
## Year2007           0.3530      0.2514     1.40    0.1645
## Year2008          -0.0841      0.2367    -0.36    0.7232
## Year2009           0.2506      0.3884     0.65    0.5209
## Year2010           0.1365      0.1139     1.20    0.2349
## Year2011           0.0400      0.1666     0.24    0.8111
## Year2012          -0.1829      0.1341    -1.36    0.1767
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.294, Adjusted R-squared:  0.0816
## Convergence in 19 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.187  0.912  0.962   0.919  0.987   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      1.04e-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 86.704 1          9.311
## LastAuthorFemale  2.575 1          1.605
## Year             204.309 16          1.181

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.36870 -0.26247  0.00373  0.24696  1.17454
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.0655     0.0138   77.25 < 2e-16 ***
## FirstAuthorFemale1  0.0425     0.1190    0.36  0.72163
## LastAuthorFemale1 -0.0571     0.1772   -0.32  0.74800
## Year1997           0.2007     0.1917    1.05  0.29840
## Year1998           0.1979     0.2634    0.75  0.45465
## Year1999           0.3530     0.1198    2.95  0.00424 **
## Year2000           0.0462     0.2696    0.17  0.86441
## Year2001           0.8216     0.1995    4.12  9.5e-05 ***
## Year2002           0.2293     0.5185    0.44  0.65955
## Year2003           0.3695     0.1951    1.89  0.06201 .
## Year2004           0.2607     0.1477    1.76  0.08160 .
## Year2005           0.0700     0.0504    1.39  0.16854
```

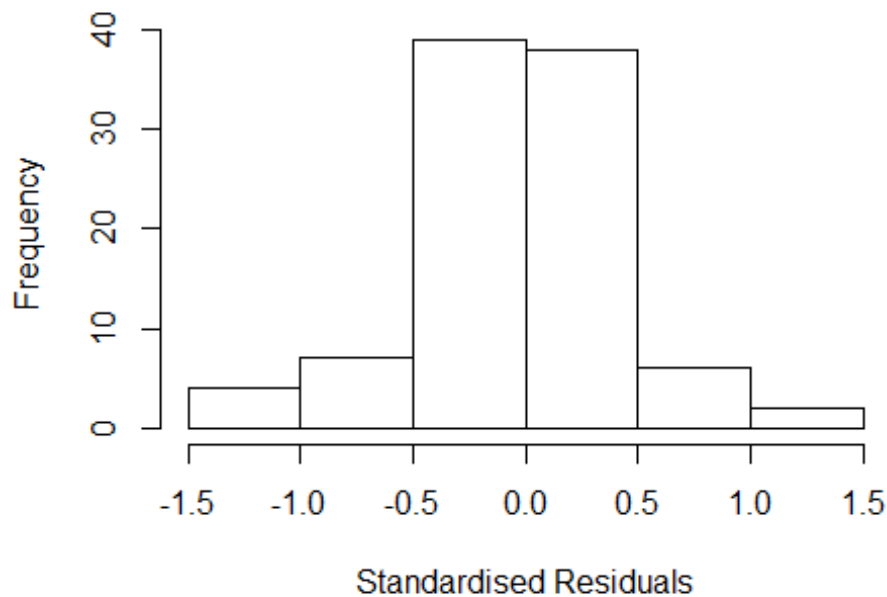


```

## Year2006          0.1810      0.1364      1.33  0.18844
## Year2007          0.4706      0.1752      2.69  0.00884 **
## Year2008          0.1731      0.2786      0.62  0.53625
## Year2009          0.2834      0.4840      0.59  0.55987
## Year2010          0.3055      0.0855      3.57  0.00061 ***
## Year2011          0.2907      0.1154      2.52  0.01381 *
## Year2012         -0.0997      0.1193     -0.84  0.40572
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.185, Adjusted R-squared:  -0.00541
## Convergence in 23 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 84 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.337  0.907   0.962   0.903   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      1.04e-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 83.02 1         9.111
## Year              83.02 16         1.148

```

Residuals from first author



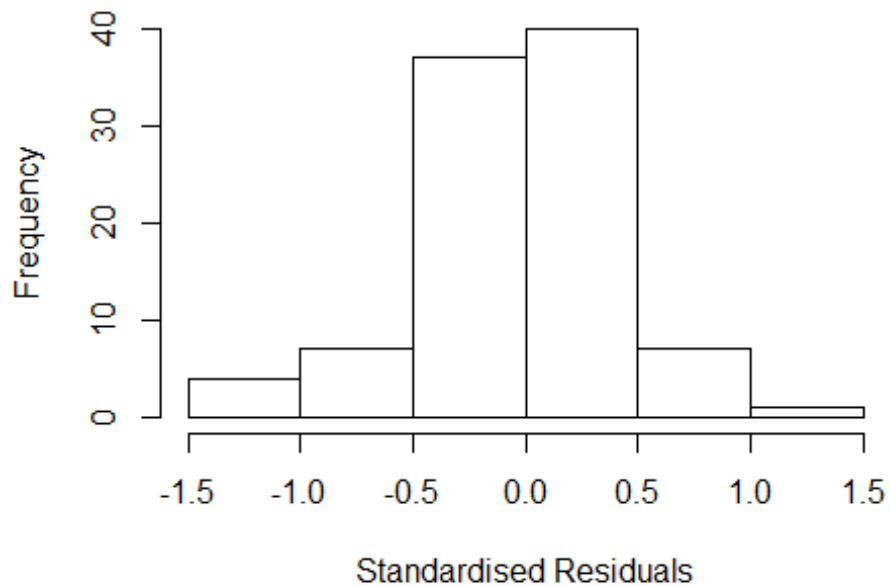
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3578 -0.2575 -0.0133 0.2444 1.2082
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0655 0.0138 77.25 < 2e-16 ***
## FirstAuthorFemale1 0.0316 0.1164 0.27 0.78659
## Year1997 0.2022 0.1910 1.06 0.29316
## Year1998 0.1912 0.2587 0.74 0.46220
## Year1999 0.3639 0.1172 3.10 0.00266 **
## Year2000 0.0285 0.2520 0.11 0.91010
## Year2001 0.8216 0.2001 4.11 9.8e-05 ***
## Year2002 0.2461 0.5041 0.49 0.62674
## Year2003 0.3728 0.1968 1.89 0.06185 .
## Year2004 0.2353 0.1097 2.15 0.03500 *
## Year2005 0.0700 0.0504 1.39 0.16852
## Year2006 0.1846 0.1351 1.37 0.17576
```

```

## Year2007          0.4770      0.1754      2.72  0.00805 **
## Year2008          0.1661      0.2781      0.60  0.55197
## Year2009          0.2607      0.5186      0.50  0.61663
## Year2010          0.2964      0.0862      3.44  0.00094 ***
## Year2011          0.2925      0.1155      2.53  0.01334 *
## Year2012         -0.0964      0.1193     -0.81  0.42161
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.443
## Multiple R-squared:  0.186, Adjusted R-squared:  0.00906
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 84 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.328  0.910   0.961   0.899   0.990   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      1.04e-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 2.258 1          1.503
## Year              2.258 16          1.026

```

Residuals from last author



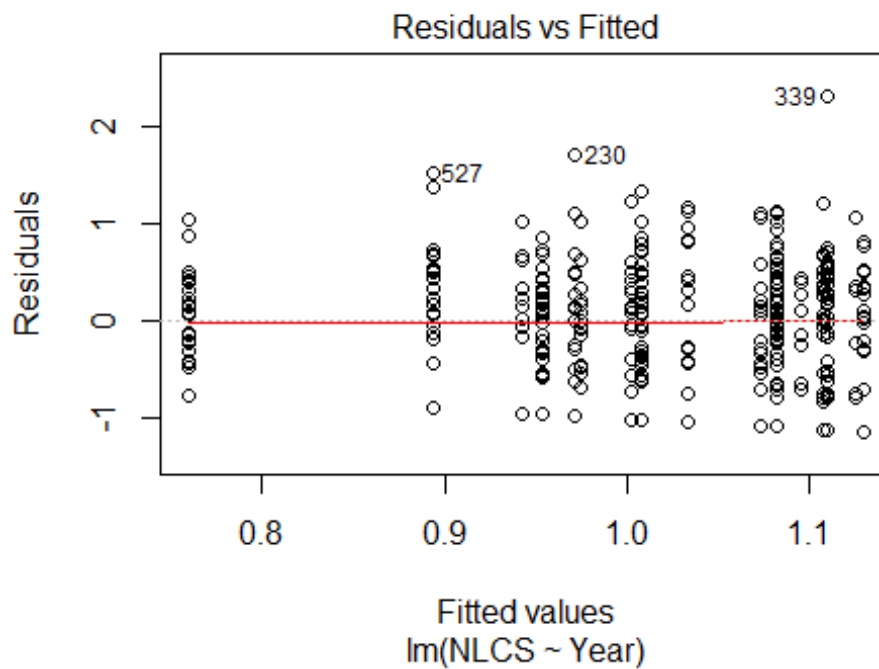
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.338288 -0.263233 -0.000268 0.243480 1.216056
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0655 0.0138 77.25 < 2e-16 ***
## LastAuthorFemale1 -0.0401 0.1718 -0.23 0.81604
## Year1997 0.2088 0.1859 1.12 0.26465
## Year1998 0.2034 0.2764 0.74 0.46387
## Year1999 0.3955 0.0138 28.67 < 2e-16 ***
## Year2000 0.0412 0.2656 0.16 0.87716
## Year2001 0.8216 0.2000 4.11 9.8e-05 ***
## Year2002 0.2430 0.5107 0.48 0.63558
## Year2003 0.3814 0.1962 1.94 0.05546 .
## Year2004 0.2728 0.1510 1.81 0.07463 .
## Year2005 0.0700 0.0504 1.39 0.16852
## Year2006 0.1951 0.1253 1.56 0.12357
```

```

## Year2007          0.4960      0.1617      3.07  0.00297 **
## Year2008          0.1823      0.2766      0.66  0.51180
## Year2009          0.2844      0.4945      0.58  0.56677
## Year2010          0.3121      0.0807      3.87  0.00023 ***
## Year2011          0.2975      0.1129      2.63  0.01016 *
## Year2012         -0.0858      0.1131     -0.76  0.45046
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared:  0.185, Adjusted R-squared:  0.00793
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 86 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.345  0.909   0.965   0.902   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      1.04e-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: 96"
## [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
##   13   24   29   16   30   18   16   27   19   19   31   30   35   48   42
## 2011 2012
##   50   40
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12   18   20    6   18   16   10   25   15   17   24   22   31   39   35
## 2011 2012

```

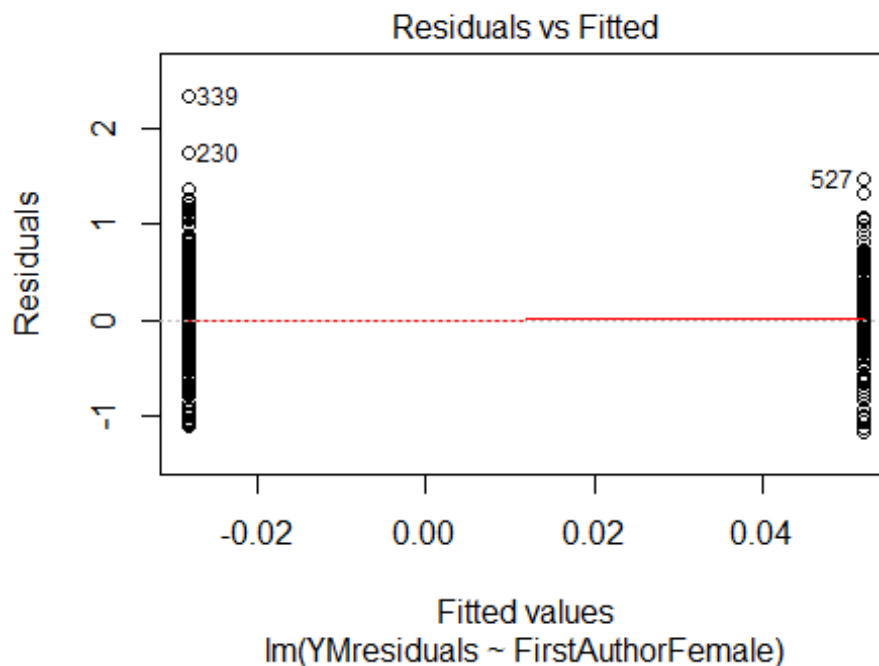
```
## 38 31
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 15 20 6 18 15 9 24 15 16 24 21 30 39 33
## 2011 2012
## 37 30
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: 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 = 0.027, df = 1, p-value = 0.9
## [1] "Female first author team size 2018 geometric mean: 1.54449937352365"
## [1] "Male first author team size 2018 geometric mean: 1.11758280649759"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

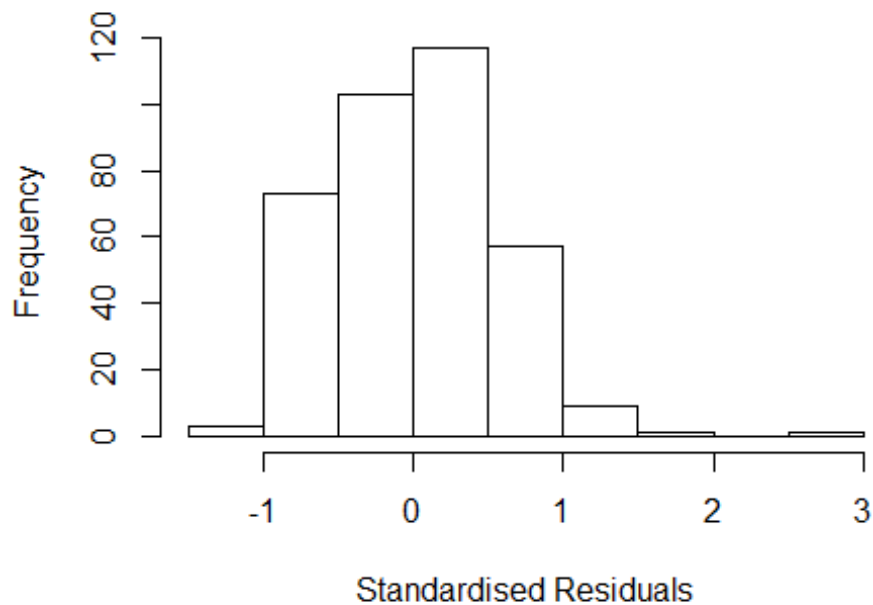
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 350, p-value = 0.03
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.3677837282176"
## [1] "Male last author team size 2018 geometric mean: 1.24200807524092"

## 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 = 300, 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 2.151  1      1.467
## LastAuthorFemale  2.175  1      1.475
## UniqueAuthors    2.375  4      1.114
## Year              3.460 16      1.040
```

Residuals from first and last author and team size



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 339 38849185241 3.414 2008      1200      2      2.55
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.4464 -0.3900  0.0126  0.3800  2.5502
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.87560    0.20318   4.31 2.1e-05 ***
## FirstAuthorFemale1 0.07179    0.08337   0.86  0.390
## LastAuthorFemale1 -0.03527    0.08380  -0.42  0.674
## UniqueAuthors2    0.35368    0.08219   4.30 2.2e-05 ***
## UniqueAuthors3    0.43804    0.10160   4.31 2.1e-05 ***
## UniqueAuthors4    0.56654    0.18198   3.11  0.002 **
## UniqueAuthors5    0.29875    0.21364   1.40  0.163
## Year1997          0.11298    0.23569   0.48  0.632
## Year1998          0.09845    0.23131   0.43  0.671
## Year1999          0.09003    0.34904   0.26  0.797
```

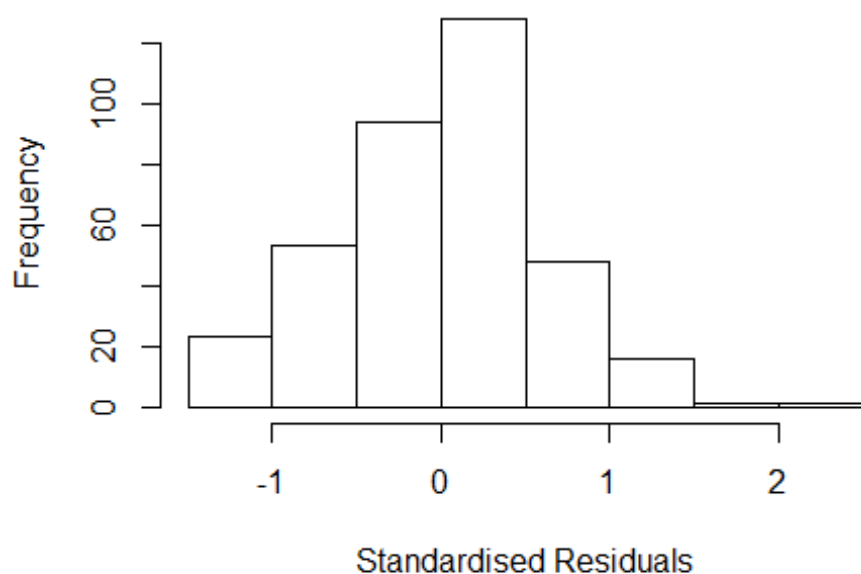


```

## Year2000      -0.00548    0.31849   -0.02    0.986
## Year2001      -0.05356    0.23846   -0.22    0.822
## Year2002       0.15181    0.25815    0.59    0.557
## Year2003      -0.23707    0.20965   -1.13    0.259
## Year2004      -0.05821    0.23054   -0.25    0.801
## Year2005      -0.06757    0.27607   -0.24    0.807
## Year2006       0.03491    0.22262    0.16    0.876
## Year2007       0.03648    0.26305    0.14    0.890
## Year2008      -0.01176    0.22391   -0.05    0.958
## Year2009      -0.00304    0.21237   -0.01    0.989
## Year2010      -0.06634    0.20689   -0.32    0.749
## Year2011       0.04687    0.21356    0.22    0.826
## Year2012      -0.14758    0.23473   -0.63    0.530
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.56
## Multiple R-squared:  0.132, Adjusted R-squared:  0.0763
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 343 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0031 0.8730 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      2.75e-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 2.113 1      1.454
## LastAuthorFemale 2.088 1      1.445
## Year      1.728 16      1.017

```

Residuals from first and last author



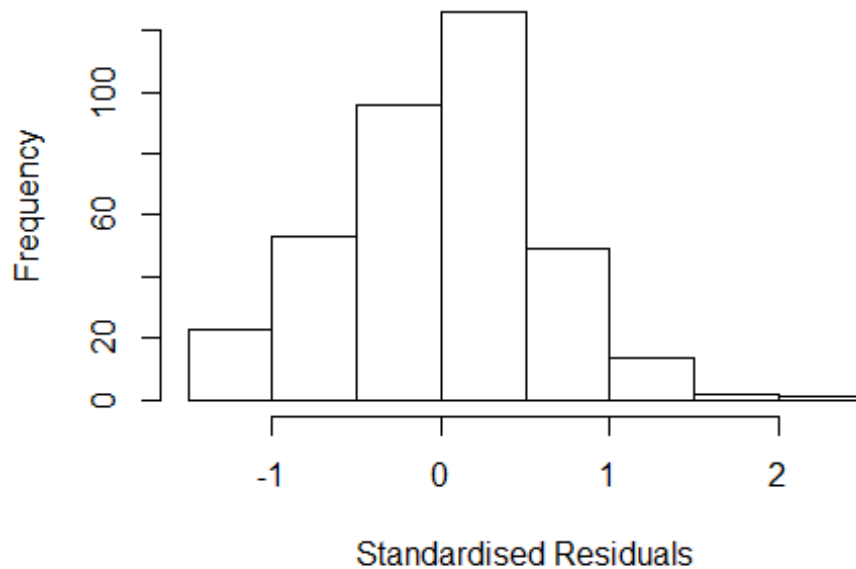
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1723 -0.3983 0.0472 0.3777 2.4118
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.91016 0.21964 4.14 4.3e-05 ***
## FirstAuthorFemale1 0.12925 0.09041 1.43 0.15
## LastAuthorFemale1 -0.03053 0.08935 -0.34 0.73
## Year1997 0.25016 0.26079 0.96 0.34
## Year1998 0.13299 0.25005 0.53 0.60
## Year1999 0.12822 0.36831 0.35 0.73
## Year2000 0.10707 0.32579 0.33 0.74
## Year2001 0.01779 0.26016 0.07 0.95
## Year2002 0.20097 0.26600 0.76 0.45
## Year2003 -0.21036 0.23214 -0.91 0.37
## Year2004 0.02148 0.23962 0.09 0.93
## Year2005 -0.10921 0.28308 -0.39 0.70
```

```

## Year2006          0.16638    0.23112    0.72    0.47
## Year2007          0.16344    0.25261    0.65    0.52
## Year2008          0.09204    0.24595    0.37    0.71
## Year2009          0.03408    0.22812    0.15    0.88
## Year2010          0.00125    0.22534    0.01    1.00
## Year2011          0.11051    0.23375    0.47    0.64
## Year2012          -0.10596    0.25112   -0.42    0.67
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.58
## Multiple R-squared:  0.0429, Adjusted R-squared:  -0.00705
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 30 weights are ~= 1. The remaining 334 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.045  0.865   0.950   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      2.75e-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 1.372 1      1.171
## Year              1.372 16      1.010

```

Residuals from first author



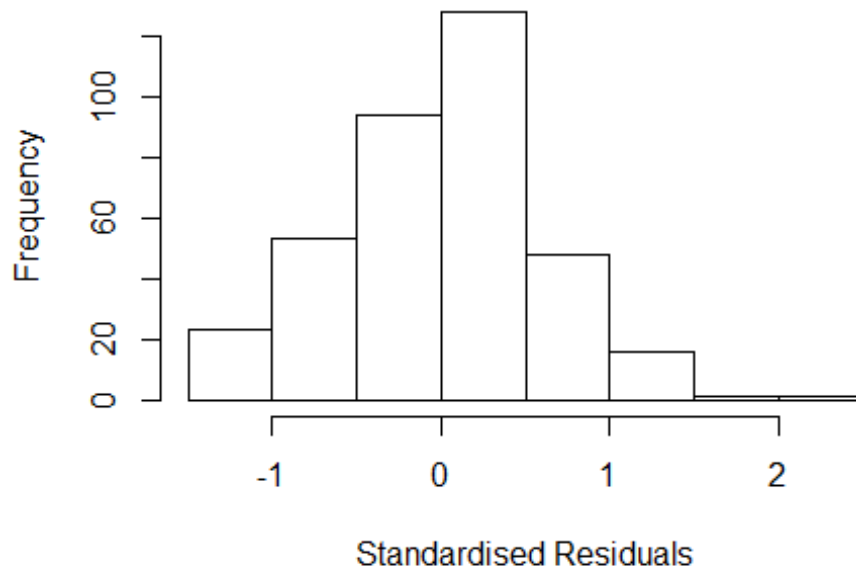
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1786 -0.3945 0.0459 0.3788 2.4168
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.90526 0.21827 4.15 4.2e-05 ***
## FirstAuthorFemale1 0.10957 0.07313 1.50 0.14
## Year1997 0.26021 0.25782 1.01 0.31
## Year1998 0.13467 0.24989 0.54 0.59
## Year1999 0.13347 0.36902 0.36 0.72
## Year2000 0.11124 0.32701 0.34 0.73
## Year2001 0.02115 0.25963 0.08 0.94
## Year2002 0.20829 0.26359 0.79 0.43
## Year2003 -0.20754 0.23114 -0.90 0.37
## Year2004 0.01553 0.23833 0.07 0.95
## Year2005 -0.10917 0.28136 -0.39 0.70
## Year2006 0.16507 0.23021 0.72 0.47
```

```

## Year2007          0.16382      0.25201      0.65      0.52
## Year2008          0.09195      0.24583      0.37      0.71
## Year2009          0.03464      0.22786      0.15      0.88
## Year2010          0.00435      0.22428      0.02      0.98
## Year2011          0.11205      0.23333      0.48      0.63
## Year2012         -0.10709      0.25056     -0.43      0.67
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.578
## Multiple R-squared:  0.0428, Adjusted R-squared:  -0.0042
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 333 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0409 0.8660 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      2.75e-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 1.355 1      1.164
## Year      1.355 16      1.010

```

Residuals from last author



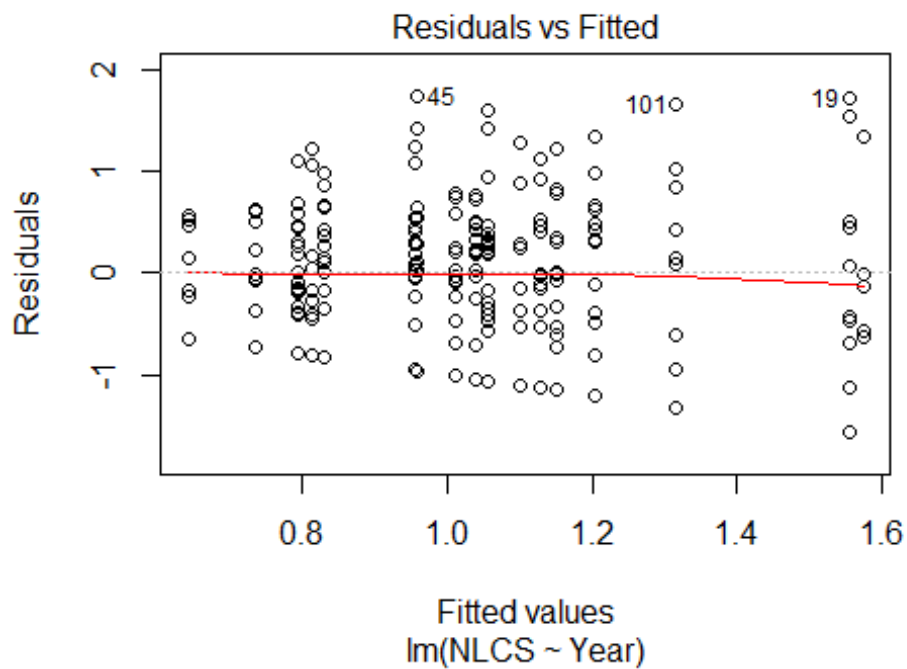
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2036 -0.3983 0.0648 0.3780 2.4018
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.93763 0.21349 4.39 1.5e-05 ***
## LastAuthorFemale1 0.05493 0.07171 0.77 0.44
## Year1997 0.26598 0.26208 1.01 0.31
## Year1998 0.10837 0.24435 0.44 0.66
## Year1999 0.14563 0.37586 0.39 0.70
## Year2000 0.08562 0.32361 0.26 0.79
## Year2001 -0.00358 0.25433 -0.01 0.99
## Year2002 0.18851 0.25926 0.73 0.47
## Year2003 -0.22950 0.22817 -1.01 0.32
## Year2004 -0.00781 0.23327 -0.03 0.97
## Year2005 -0.10935 0.27920 -0.39 0.70
## Year2006 0.13479 0.22384 0.60 0.55
```

```

## Year2007      0.15408      0.24976      0.62      0.54
## Year2008      0.07459      0.24129      0.31      0.76
## Year2009      0.03292      0.22605      0.15      0.88
## Year2010     -0.00356      0.22297     -0.02      0.99
## Year2011      0.10665      0.23171      0.46      0.65
## Year2012     -0.10894      0.24817     -0.44      0.66
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.585
## Multiple R-squared:  0.0373, Adjusted R-squared:  -0.00998
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 340 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0538 0.8720 0.9530 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      2.75e-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: 364"
## [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
##   10   15   14   17   10   21   11   16   25   16   17   20   22   31   25
## 2011 2012
##   17   26
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    8   10   10   12    5   17   10   14   18   13   10   13   17   20   19
## 2011 2012

```

```
##      9      23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      8      9      10      12      5      16      10      13      17      13      6      11      17      19      18
## 2011 2012
##      8      21
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 19, df = 16, p-value = 0.2
```

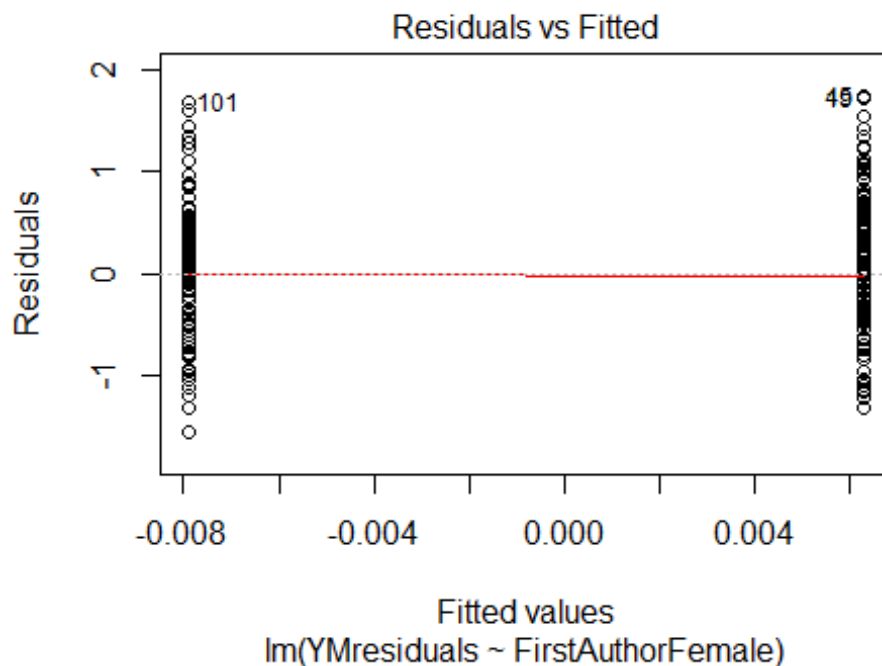


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.026, df = 1, p-value = 0.9
## [1] "Female first author team size 2018 geometric mean: 1.56508458007329"
## [1] "Male first author team size 2018 geometric mean: 1.57461010625845"
## 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 = 14, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.44224957030741"
## [1] "Male last author team size 2018 geometric mean: 1.62238960361098"

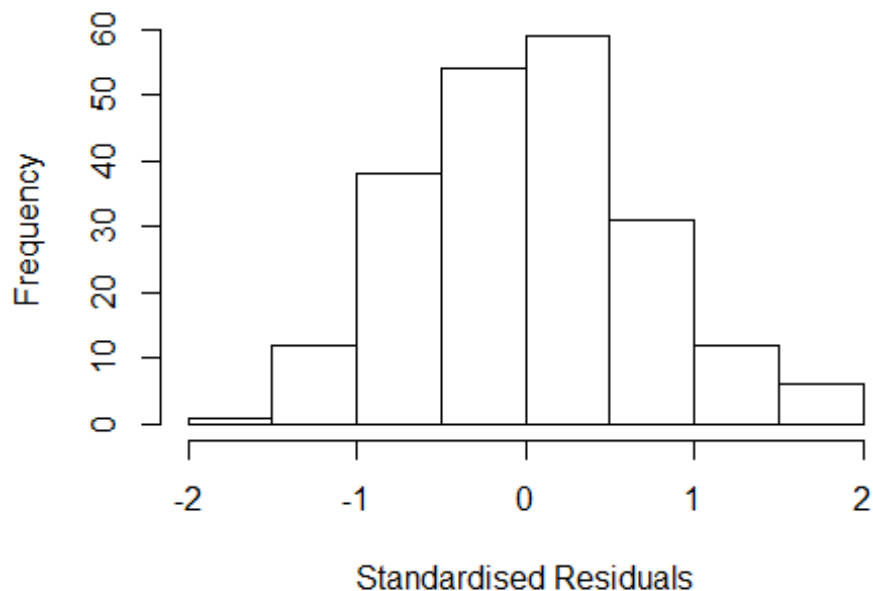
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 10, p-value = 0.8
## 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	5.212	1	2.283
LastAuthorFemale	3.822	1	1.955
UniqueAuthors	41.373	4	1.593
Year	62.055	16	1.138

Residuals from first and last author and team size



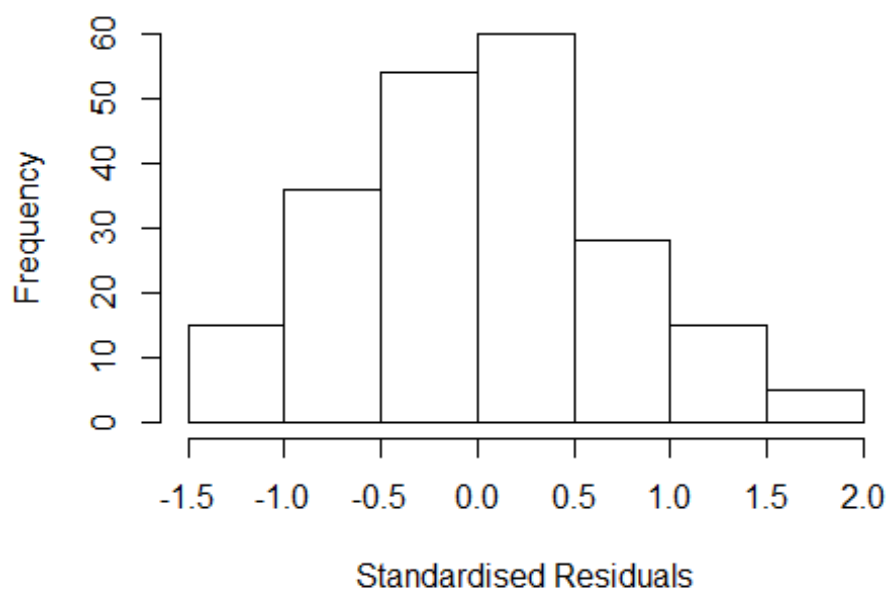
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.66001 -0.48970 0.00346 0.47120 1.95948
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6052 0.2138 2.83 0.0051 **
## FirstAuthorFemale1 0.1665 0.1351 1.23 0.2196
## LastAuthorFemale1 -0.2734 0.1326 -2.06 0.0405 *
## UniqueAuthors2 0.2228 0.1203 1.85 0.0654 .
## UniqueAuthors3 0.4114 0.1560 2.64 0.0090 **
## UniqueAuthors4 0.8050 0.3648 2.21 0.0285 *
## UniqueAuthors5 -0.0622 0.1928 -0.32 0.7474
## Year1997 0.7053 0.4055 1.74 0.0836 .
## Year1998 0.1567 0.2556 0.61 0.5407
## Year1999 0.2388 0.3517 0.68 0.4981
```

```

## Year2000          0.8345      0.2651      3.15      0.0019 **
## Year2001          0.5320      0.3036      1.75      0.0814 .
## Year2002          0.6655      0.4964      1.34      0.1816
## Year2003          0.4560      0.2881      1.58      0.1151
## Year2004          0.3185      0.2457      1.30      0.1964
## Year2005          0.4747      0.2889      1.64      0.1020
## Year2006         -0.1186      0.2656     -0.45      0.6559
## Year2007          0.3405      0.2550      1.33      0.1835
## Year2008          0.3943      0.2477      1.59      0.1130
## Year2009          0.1793      0.2406      0.75      0.4572
## Year2010          0.2140      0.2343      0.91      0.3622
## Year2011          0.2366      0.3473      0.68      0.4965
## Year2012          0.2586      0.2442      1.06      0.2910
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.739
## Multiple R-squared:  0.131, Adjusted R-squared:  0.0306
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 19 weights are ~= 1. The remaining 194 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.462  0.888  0.956  0.920  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      4.69e-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 1.885 1      1.373
## LastAuthorFemale  1.835 1      1.355
## Year              2.196 16      1.025

```

Residuals from first and last author



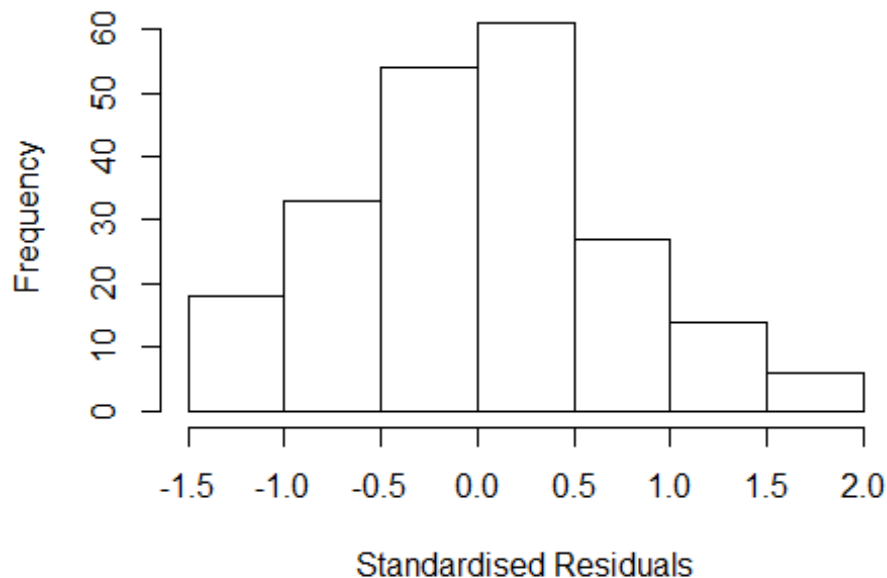
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.42722 -0.46652 0.00963 0.44762 1.85005
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6905 0.1999 3.45 0.00068 ***
## FirstAuthorFemale1 0.1110 0.1293 0.86 0.39185
## LastAuthorFemale1 -0.2022 0.1235 -1.64 0.10336
## Year1997 0.7294 0.4360 1.67 0.09598 .
## Year1998 0.1193 0.2631 0.45 0.65078
## Year1999 0.2030 0.3691 0.55 0.58295
## Year2000 0.8758 0.3670 2.39 0.01799 *
## Year2001 0.5147 0.3004 1.71 0.08825 .
## Year2002 0.6257 0.4582 1.37 0.17369
## Year2003 0.4560 0.2789 1.63 0.10370
## Year2004 0.2954 0.2437 1.21 0.22694
## Year2005 0.4920 0.2686 1.83 0.06856 .
```

```

## Year2006          -0.0358      0.2625    -0.14  0.89156
## Year2007           0.3604      0.2499     1.44  0.15075
## Year2008           0.4439      0.2321     1.91  0.05726 .
## Year2009           0.1514      0.2276     0.67  0.50655
## Year2010           0.1580      0.2279     0.69  0.48896
## Year2011           0.3041      0.3616     0.84  0.40145
## Year2012           0.3980      0.2426     1.64  0.10256
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.73
## Multiple R-squared:  0.0825, Adjusted R-squared:  -0.00259
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 18 weights are ~= 1. The remaining 195 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.500  0.869   0.957   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      4.69e-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 1.459 1      1.208
## Year              1.459 16      1.012

```

Residuals from first author

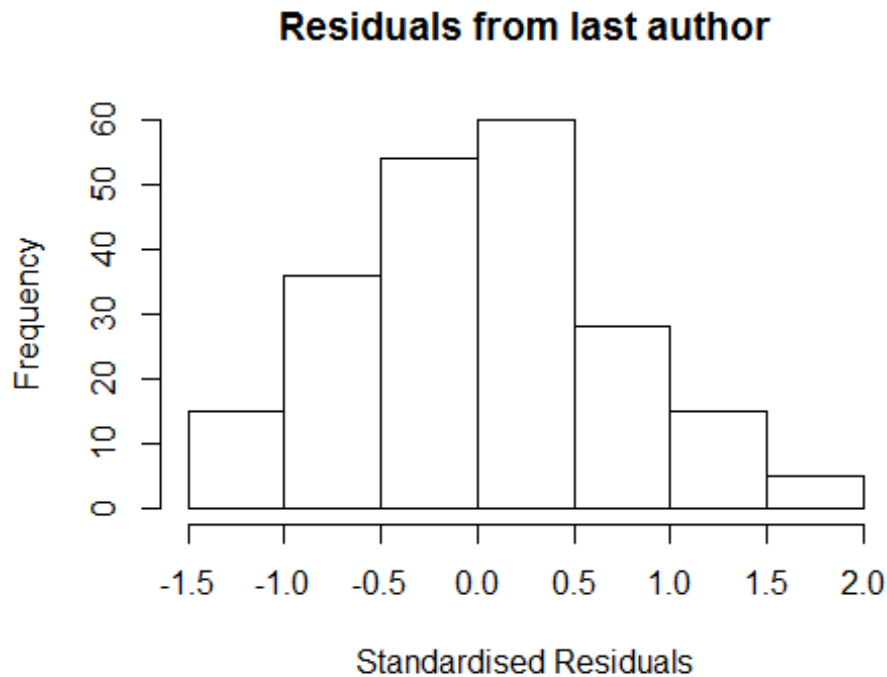


```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3626 -0.4748 0.0272 0.4617 1.8902
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6587 0.1968 3.35 0.00098 ***
## FirstAuthorFemale1 -0.0172 0.1166 -0.15 0.88255
## Year1997 0.7211 0.4691 1.54 0.12584
## Year1998 0.1200 0.2683 0.45 0.65528
## Year1999 0.2254 0.3880 0.58 0.56206
## Year2000 0.8404 0.3641 2.31 0.02205 *
## Year2001 0.5206 0.2954 1.76 0.07957 .
## Year2002 0.6526 0.4329 1.51 0.13330
## Year2003 0.4867 0.2723 1.79 0.07546 .
## Year2004 0.2846 0.2444 1.16 0.24565
## Year2005 0.4794 0.2725 1.76 0.08002 .
## Year2006 0.0194 0.2454 0.08 0.93705
```

```

## Year2007          0.3653      0.2487      1.47  0.14350
## Year2008          0.4175      0.2316      1.80  0.07302 .
## Year2009          0.1513      0.2241      0.68  0.50035
## Year2010          0.1360      0.2297      0.59  0.55447
## Year2011          0.3809      0.3392      1.12  0.26285
## Year2012          0.3878      0.2487      1.56  0.12053
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.738
## Multiple R-squared:  0.0716, Adjusted R-squared:  -0.0093
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 191 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.491  0.871  0.954  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      4.69e-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 1.382 1      1.176
## Year              1.382 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
## -1.3422 -0.4658 0.0282 0.4676 1.8328
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.7324 0.1971 3.72 0.00026 ***
## LastAuthorFemale1 -0.1362 0.1102 -1.24 0.21795
## Year1997 0.7048 0.4409 1.60 0.11158
## Year1998 0.0962 0.2626 0.37 0.71458
## Year1999 0.1688 0.3737 0.45 0.65206
## Year2000 0.8257 0.3568 2.31 0.02170 *
## Year2001 0.4816 0.2985 1.61 0.10828
## Year2002 0.6098 0.4412 1.38 0.16851
## Year2003 0.4323 0.2853 1.52 0.13129
## Year2004 0.2796 0.2466 1.13 0.25824
## Year2005 0.4426 0.2648 1.67 0.09626 .
## Year2006 -0.0226 0.2692 -0.08 0.93332
```

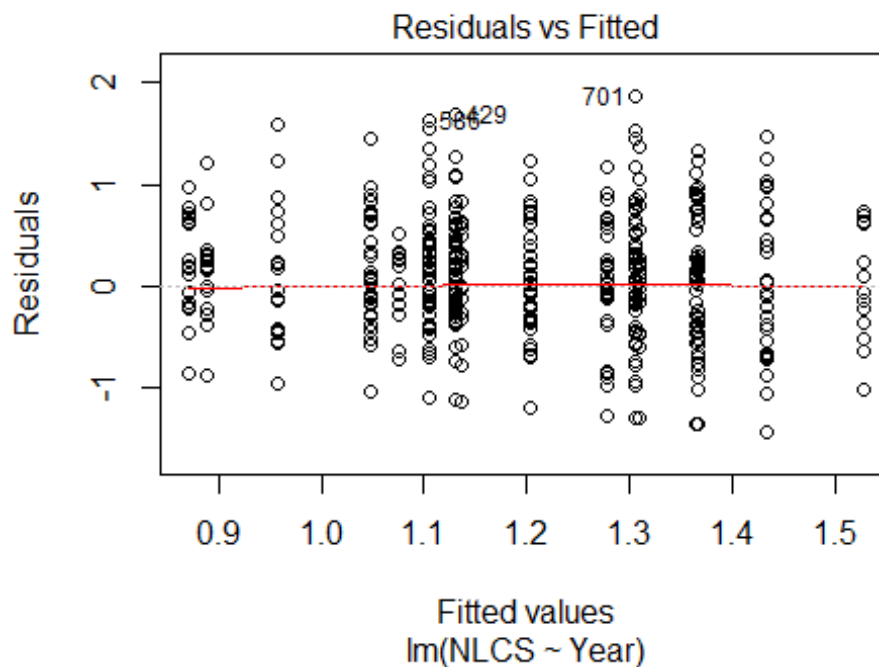


```

## Year2007          0.3361      0.2495      1.35  0.17944
## Year2008          0.4248      0.2360      1.80  0.07344 .
## Year2009          0.1412      0.2314      0.61  0.54232
## Year2010          0.1154      0.2257      0.51  0.60977
## Year2011          0.3280      0.3643      0.90  0.36901
## Year2012          0.3711      0.2479      1.50  0.13594
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.722
## Multiple R-squared:  0.0795, Adjusted R-squared:  -0.000719
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 189 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.499  0.863   0.952   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      4.69e-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: 213"
## [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
##   12   27   27   25   17   24   24   25   28   39   44   49   45   59   43
## 2011 2012
##   76   73
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   12   21   22   21   13   23   21   22   21   31   36   42   42   45   33
## 2011 2012

```

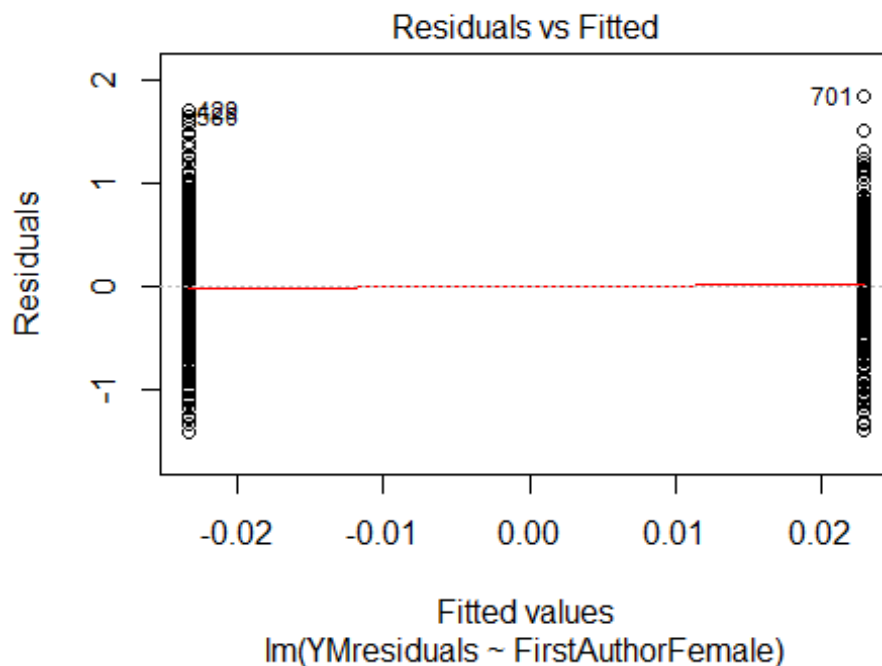
```
## 57 58
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 12 21 22 21 12 21 21 22 20 29 36 41 42 45 33
## 2011 2012
## 55 54
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 12, df = 16, p-value = 0.8
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.2, df = 1, p-value = 0.07
## [1] "Female first author team size 2018 geometric mean: 1.47779187693649"
## [1] "Male first author team size 2018 geometric mean: 1.23994462681838"
## 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 = 300, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.54099921734491"
## [1] "Male last author team size 2018 geometric mean: 1.11849604597382"

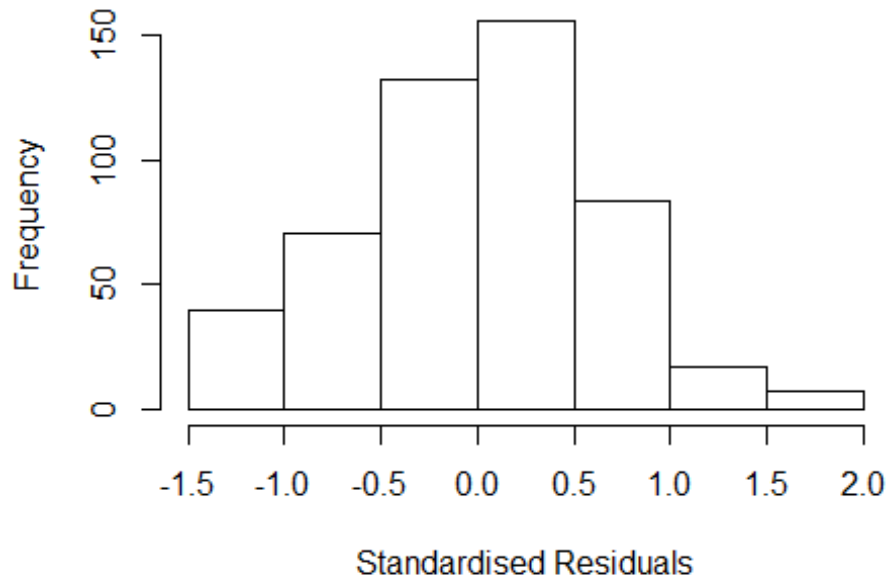
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 320, 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	2.708	1	1.646
LastAuthorFemale	2.619	1	1.618
UniqueAuthors	1.799	4	1.076
Year	1.943	16	1.021

Residuals from first and last author and team size



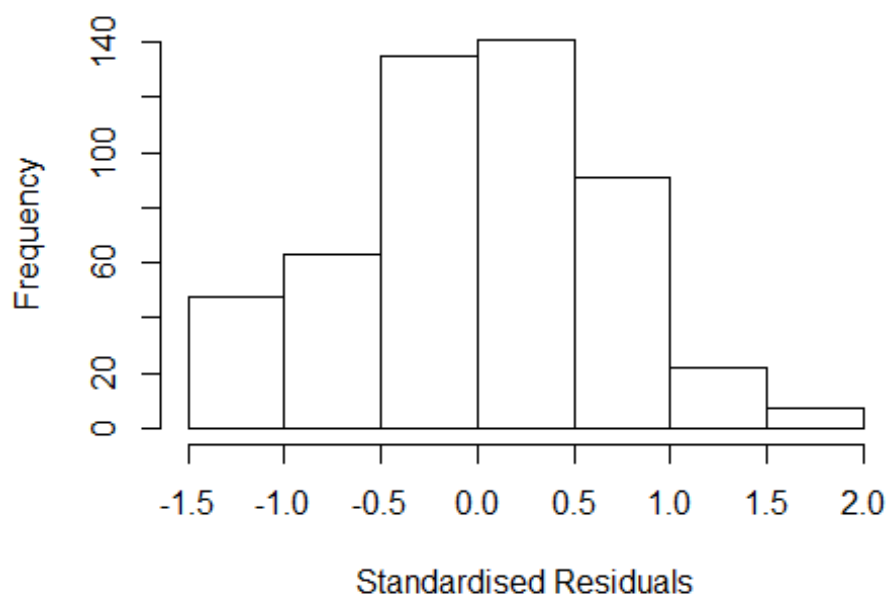
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4899 -0.4251 0.0208 0.4221 1.9810
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.97496 0.11557 8.44 3.8e-16 ***
## FirstAuthorFemale1 0.00218 0.09400 0.02 0.98154
## LastAuthorFemale1 0.01941 0.09218 0.21 0.83329
## UniqueAuthors2 0.27899 0.07794 3.58 0.00038 ***
## UniqueAuthors3 0.38322 0.11320 3.39 0.00077 ***
## UniqueAuthors4 0.24866 0.19016 1.31 0.19161
## UniqueAuthors5 -0.11454 0.19319 -0.59 0.55354
## Year1997 0.13399 0.17289 0.77 0.43873
## Year1998 -0.12919 0.17012 -0.76 0.44800
## Year1999 -0.16824 0.17389 -0.97 0.33376
```

```

## Year2000          0.53378    0.20717    2.58  0.01027 *
## Year2001         -0.11605    0.18723   -0.62  0.53567
## Year2002          0.30368    0.18672    1.63  0.10452
## Year2003          0.10977    0.16992    0.65  0.51858
## Year2004          0.24688    0.21239    1.16  0.24565
## Year2005          0.43208    0.19034    2.27  0.02364 *
## Year2006          0.24385    0.16819    1.45  0.14776
## Year2007          0.08678    0.13293    0.65  0.51420
## Year2008          0.06978    0.14142    0.49  0.62195
## Year2009         -0.01066    0.14586   -0.07  0.94180
## Year2010          0.21728    0.15178    1.43  0.15293
## Year2011          0.04377    0.14474    0.30  0.76245
## Year2012          0.19546    0.14333    1.36  0.17330
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.666
## Multiple R-squared:  0.1,   Adjusted R-squared:  0.0593
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 42 weights are ~= 1. The remaining 465 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.357  0.865   0.956   0.913   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      1.97e-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 2.662 1      1.632
## LastAuthorFemale  2.652 1      1.628
## Year              1.260 16      1.007

```

Residuals from first and last author



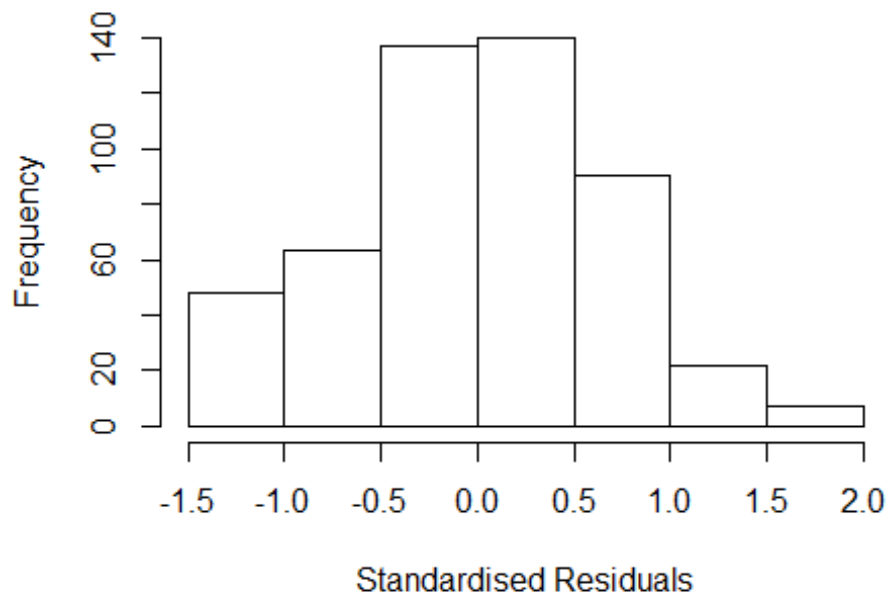
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4535 -0.3984 0.0155 0.4474 1.8693
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0432 0.1210 8.62 <2e-16 ***
## FirstAuthorFemale1 0.0296 0.0964 0.31 0.759
## LastAuthorFemale1 0.0329 0.0960 0.34 0.732
## Year1997 0.0814 0.1714 0.47 0.635
## Year1998 -0.2077 0.1702 -1.22 0.223
## Year1999 -0.1915 0.1761 -1.09 0.277
## Year2000 0.4678 0.2092 2.24 0.026 *
## Year2001 -0.1930 0.1939 -1.00 0.320
## Year2002 0.3311 0.2013 1.64 0.101
## Year2003 0.0628 0.1755 0.36 0.721
## Year2004 0.2742 0.2136 1.28 0.200
## Year2005 0.4103 0.1929 2.13 0.034 *
```

```

## Year2006          0.2871      0.1718      1.67      0.095 .
## Year2007          0.0968      0.1415      0.68      0.494
## Year2008          0.0669      0.1497      0.45      0.655
## Year2009         -0.0168      0.1518     -0.11      0.912
## Year2010          0.2183      0.1598      1.37      0.172
## Year2011          0.0401      0.1489      0.27      0.788
## Year2012          0.1980      0.1497      1.32      0.187
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.673
## Multiple R-squared:  0.0648, Adjusted R-squared:  0.0303
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 47 weights are ~= 1. The remaining 460 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.421  0.865   0.949   0.910   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      1.97e-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 1.111 1          1.054
## Year              1.111 16          1.003

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.460 -0.394 0.015 0.450 1.874
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.0447 0.1202 8.69 <2e-16 ***
## FirstAuthorFemale1 0.0564 0.0624 0.90 0.367
## Year1997 0.0819 0.1710 0.48 0.632
## Year1998 -0.2055 0.1699 -1.21 0.227
## Year1999 -0.1873 0.1755 -1.07 0.286
## Year2000 0.4707 0.2084 2.26 0.024 *
## Year2001 -0.1922 0.1938 -0.99 0.322
## Year2002 0.3313 0.2007 1.65 0.099 .
## Year2003 0.0639 0.1750 0.37 0.715
## Year2004 0.2684 0.2090 1.28 0.200
## Year2005 0.4148 0.1927 2.15 0.032 *
## Year2006 0.2929 0.1721 1.70 0.089 .
```



```

## Year2007          0.0999      0.1415      0.71      0.480
## Year2008          0.0673      0.1492      0.45      0.652
## Year2009         -0.0160      0.1514     -0.11      0.916
## Year2010          0.2209      0.1595      1.39      0.167
## Year2011          0.0418      0.1482      0.28      0.778
## Year2012          0.1976      0.1492      1.32      0.186
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.672
## Multiple R-squared:  0.0645, Adjusted R-squared:  0.032
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 45 weights are ~= 1. The remaining 462 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.416  0.863  0.951  0.910  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      1.97e-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 1.107 1      1.052
## 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.4536 -0.3988  0.0143  0.4535  1.8705

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.0490    0.1195    8.78 <2e-16 ***
## LastAuthorFemale1 0.0571    0.0623    0.92  0.360
## Year1997        0.0803    0.1720    0.47  0.641
## Year1998       -0.2102    0.1706   -1.23  0.218
## Year1999       -0.1983    0.1738   -1.14  0.254
## Year2000        0.4658    0.2093    2.23  0.026 *
## Year2001       -0.1966    0.1941   -1.01  0.312
## Year2002        0.3296    0.2021    1.63  0.104
## Year2003        0.0593    0.1754    0.34  0.736
## Year2004        0.2771    0.2130    1.30  0.194
## Year2005        0.4046    0.1922    2.10  0.036 *
## Year2006        0.2814    0.1717    1.64  0.102
## Year2007        0.0939    0.1417    0.66  0.508
## Year2008        0.0647    0.1504    0.43  0.667
## Year2009       -0.0191    0.1525   -0.13  0.900
## Year2010        0.2141    0.1593    1.34  0.180
## Year2011        0.0370    0.1492    0.25  0.804
## Year2012        0.1964    0.1503    1.31  0.192
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.673
## Multiple R-squared:  0.0647, Adjusted R-squared:  0.0322
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 47 weights are ~= 1. The remaining 460 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.420  0.864  0.947  0.910  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      1.97e-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: 507"
## [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 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    2    7    4    4    4    4    4    2    3    5    6    6    8    7
## 2012
##    9
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    0    3    2    2    1    3    4    1    3    5    6    6    8    6
## 2012
##    8
##
## 1996 1997 1998 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    1    0    3    2    1    1    2    4    1    3    5    6    6    8    6
## 2012
##    8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.32573393556655"
## [1] "Male first author team size 2018 geometric mean: 3.06683307221104"

## 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 = 33, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.16227766016838"
## [1] "Male last author team size 2018 geometric mean: 3.17970925612606"

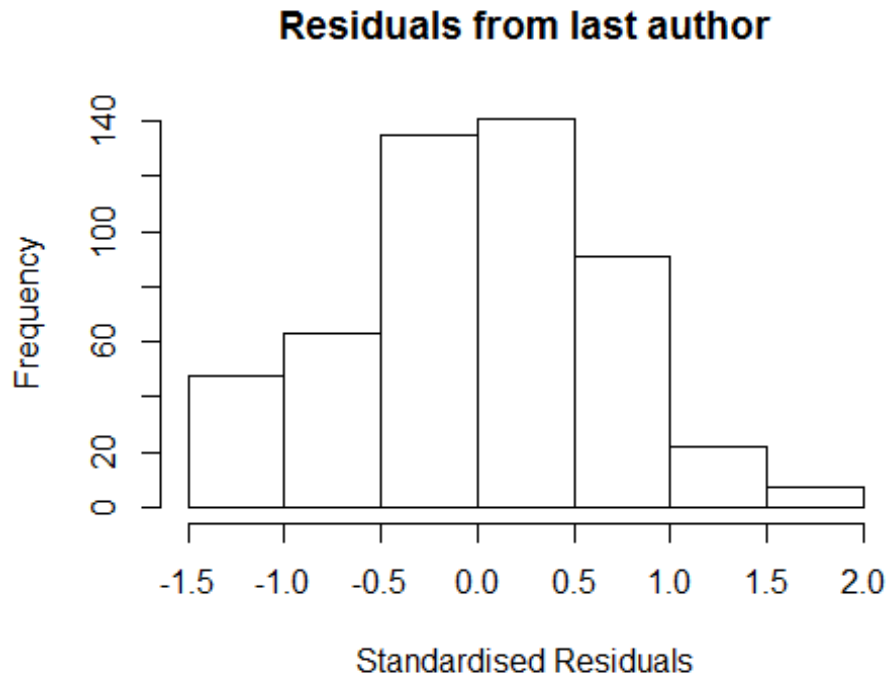
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 14, 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"

```

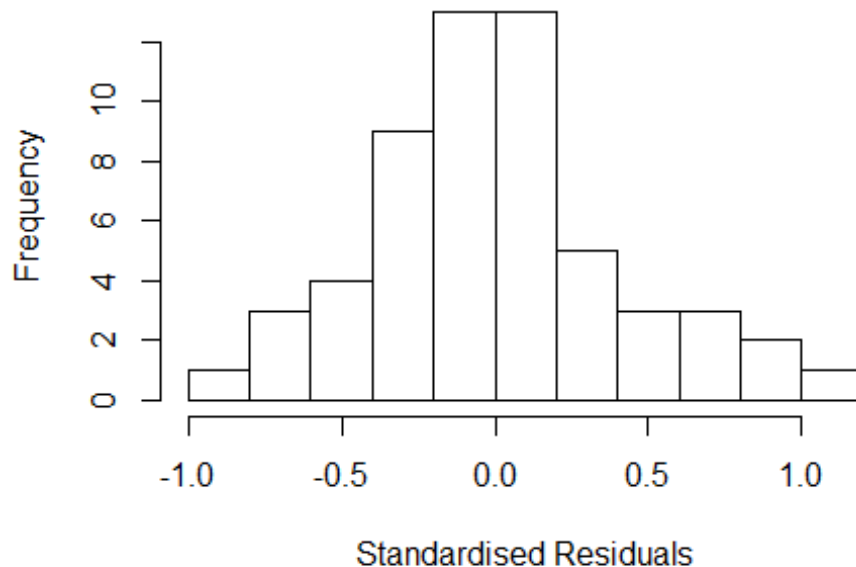
```
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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
## LastAuthorFemale	NaN	1	NaN
## UniqueAuthors	NaN	4	NaN
## Year	NaN	14	NaN

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -8.55e-01 -2.43e-01 5.00e-16 1.95e-01 1.08e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8403 0.2007 4.19 0.00017 ***
## FirstAuthorFemale1 -0.3929 0.1426 -2.76 0.00914 **
## LastAuthorFemale1 -0.2042 0.1532 -1.33 0.19095
## UniqueAuthors2 0.0297 0.1885 0.16 0.87580
## UniqueAuthors3 0.4107 0.2007 2.05 0.04810 *
## UniqueAuthors4 0.4485 0.1870 2.40 0.02175 *
## UniqueAuthors5 0.3498 0.2205 1.59 0.12130
## Year1998 -0.0657 0.3473 -0.19 0.85100
## Year2000 0.0147 1.7666 0.01 0.99340
## Year2001 0.3372 0.1499 2.25 0.03074 *
```

```

## Year2002          0.2890      0.0000      Inf < 2e-16 ***
## Year2003         -0.3005      0.3710     -0.81 0.42339
## Year2004          0.6780      0.3795      1.79 0.08239 .
## Year2005          0.3732      0.2951      1.26 0.21407
## Year2006          0.3721      0.2287      1.63 0.11234
## Year2007         -0.1355      0.4001     -0.34 0.73693
## Year2008          0.4328      0.2086      2.07 0.04523 *
## Year2009          0.4011      0.1605      2.50 0.01715 *
## Year2010          0.3319      0.2852      1.16 0.25225
## Year2011          0.2315      0.2378      0.97 0.33674
## Year2012          0.0916      0.2305      0.40 0.69359
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.486
## Multiple R-squared:  0.376, Adjusted R-squared:  0.0288
## Convergence in 32 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 48 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.604  0.886  0.968  0.926  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      1.75e-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 3.671e+14 1      1.916e+07
## LastAuthorFemale  3.664e+14 1      1.914e+07
## Year              1.082e+15 14      3.443e+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
## -0.89520 -0.23038  0.00615  0.21959  1.27485
##
## Coefficients:
##              Estimate Std. Error   t value Pr(>|t|)
## (Intercept)    1.25e+00   3.42e-08  3.66e+07  <2e-16 ***
## FirstAuthorFemale1 -3.41e-01   2.04e-01 -1.68e+00  0.1017
## LastAuthorFemale1 -1.51e-01   2.07e-01 -7.30e-01  0.4699
## Year1998        -4.80e-01   2.83e-01 -1.70e+00  0.0977 .
## Year2000        -3.96e-01   2.45e+00 -1.60e-01  0.8727
## Year2001         2.32e-01   1.54e-01  1.51e+00  0.1399
## Year2002         2.89e-01   1.75e-08  1.65e+07  <2e-16 ***
## Year2003        -5.40e-01   1.79e-01 -3.01e+00  0.0045 **
## Year2004         2.09e-01   4.14e-01  5.10e-01  0.6162
## Year2005        -1.13e-01   1.54e-01 -7.40e-01  0.4665
## Year2006        -2.82e-02   1.21e-01 -2.30e-01  0.8179
## Year2007        -3.56e-01   2.94e-01 -1.21e+00  0.2330
## Year2008         2.57e-01   1.29e-01  1.99e+00  0.0536 .
## Year2009         1.66e-01   1.67e-01  9.90e-01  0.3260
## Year2010         1.53e-02   2.22e-01  7.00e-02  0.9455
## Year2011        -2.56e-03   2.18e-01 -1.00e-02  0.9907
## Year2012        -1.97e-01   1.97e-01 -1.00e+00  0.3254
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.458
## Multiple R-squared:  0.29,   Adjusted R-squared:  0.00558
## Convergence in 48 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 49 ones are summarized as
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      0.419  0.867  0.967  0.907  0.988  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      1.75e-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

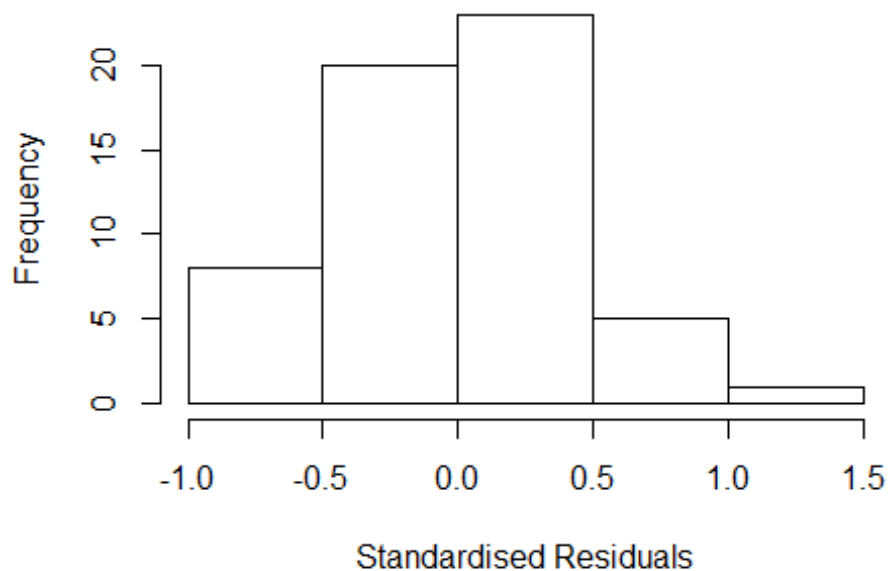
```

```
##           "bisquare"           "nonsingular"           ".vcov.avar1"
## compute.outlier.stats
##           "SM"
## seed : int(0)
## [1] "Regression 3: First author gender, Year as factors"

## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

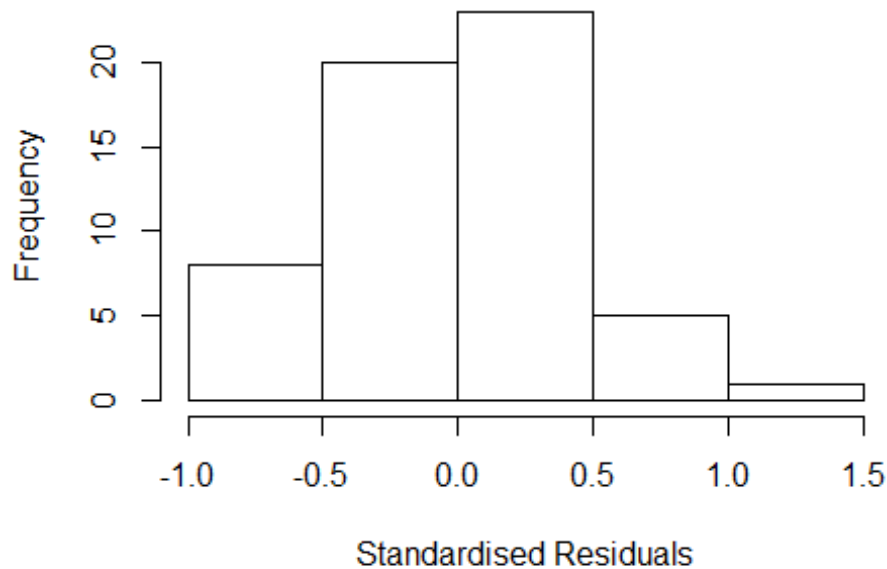
## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author



```
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1 NaN
## Year NaN 14 NaN
```


Residuals from first author



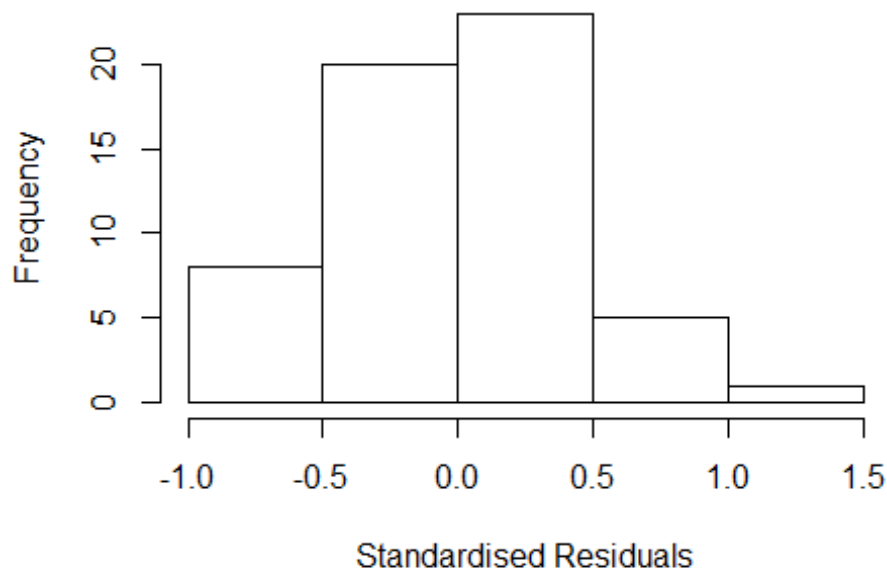
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.91633 -0.22712 0.00616 0.17477 1.45942
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.25100 0.00000 Inf <2e-16 ***
## FirstAuthorFemale1 -0.41067 0.15851 -2.59 0.0132 *
## Year1998 -0.47988 0.28215 -1.70 0.0966 .
## Year2000 -0.39600 2.37271 -0.17 0.8683
## Year2001 0.15067 0.15851 0.95 0.3474
## Year2002 0.28900 0.00000 Inf <2e-16 ***
## Year2003 -0.58017 0.21096 -2.75 0.0088 **
## Year2004 0.09425 0.35046 0.27 0.7893
## Year2005 -0.19433 0.15851 -1.23 0.2272
## Year2006 -0.02816 0.12143 -0.23 0.8178
## Year2007 -0.33467 0.29311 -1.14 0.2602
## Year2008 0.19234 0.11319 1.70 0.0969 .
```

```

## Year2009          0.17170      0.19942      0.86      0.3943
## Year2010          -0.00986      0.23260     -0.04      0.9664
## Year2011          -0.06611      0.20128     -0.33      0.7443
## Year2012          -0.24706      0.17073     -1.45      0.1555
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.461
## Multiple R-squared:  0.285, Adjusted R-squared:  0.0239
## Convergence in 46 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 46 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.294  0.860  0.966  0.899  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      1.75e-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 7.863e+13 1      8.867e+06
## Year              7.863e+13 14      3.135e+00

```

Residuals from last author



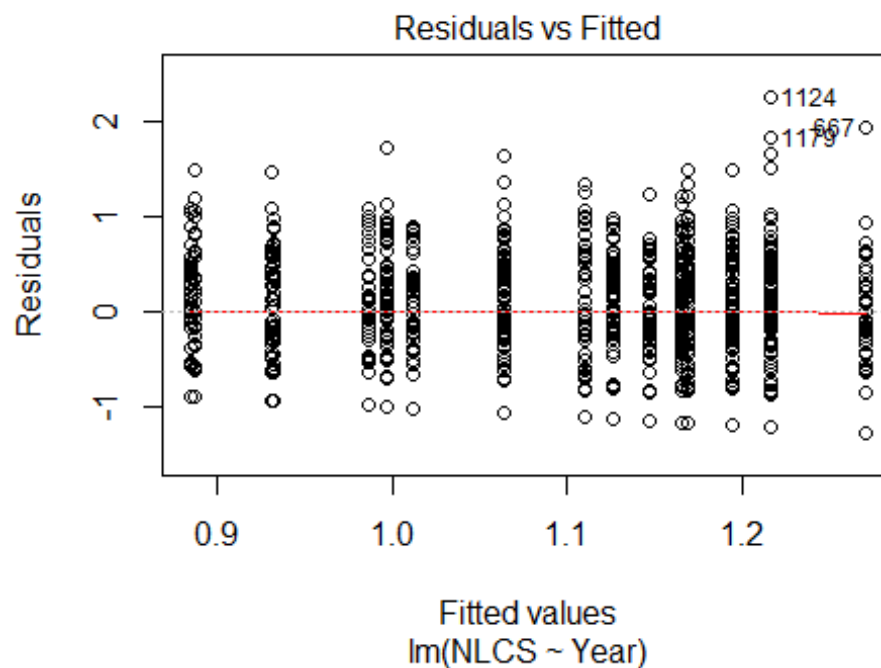
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.00388 -0.23982 0.00615 0.23090 0.99098
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.25e+00 3.64e-08 3.44e+07 <2e-16 ***
## LastAuthorFemale1 -3.14e-01 1.60e-01 -1.96e+00 0.056 .
## Year1998 -4.80e-01 2.83e-01 -1.69e+00 0.098 .
## Year2000 -3.96e-01 2.42e+00 -1.60e-01 0.871
## Year2001 5.37e-02 1.60e-01 3.40e-01 0.738
## Year2002 2.89e-01 3.78e-08 7.65e+06 <2e-16 ***
## Year2003 -6.29e-01 2.51e-01 -2.50e+00 0.016 *
## Year2004 2.58e-01 3.37e-01 7.70e-01 0.449
## Year2005 -2.91e-01 1.60e-01 -1.82e+00 0.076 .
## Year2006 -2.82e-02 1.21e-01 -2.30e-01 0.818
## Year2007 -4.61e-01 2.70e-01 -1.71e+00 0.096 .
## Year2008 1.98e-01 1.45e-01 1.36e+00 0.180
```

```

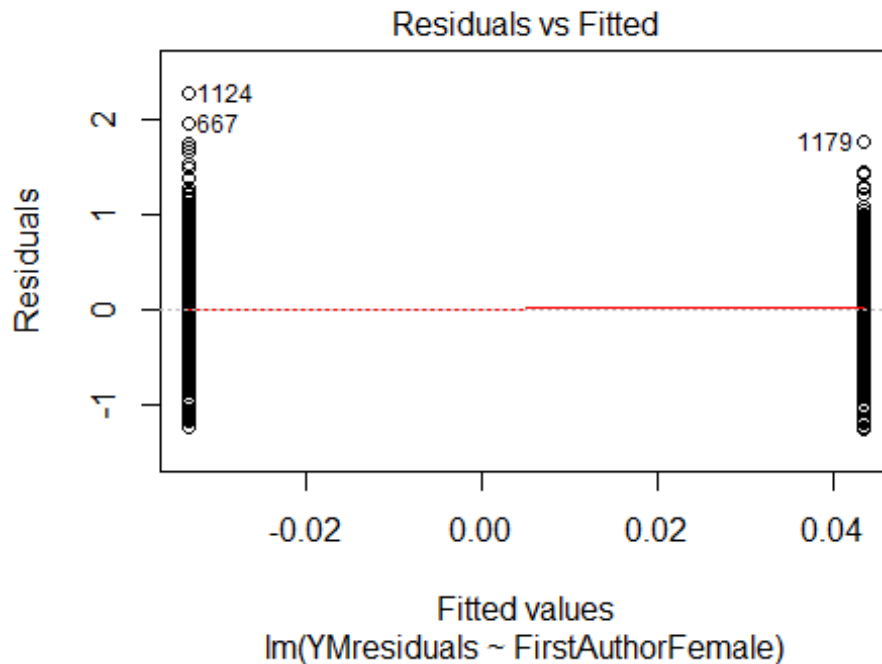
## Year2009          -1.52e-02   9.71e-02 -1.60e-01   0.877
## Year2010          -1.17e-01   1.93e-01 -6.00e-01   0.549
## Year2011          -8.81e-02   2.16e-01 -4.10e-01   0.685
## Year2012          -2.47e-01   3.24e-01 -7.60e-01   0.450
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.459
## Multiple R-squared:  0.238, Adjusted R-squared:  -0.0401
## Convergence in 47 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 47 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.612  0.826  0.959  0.897  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      1.75e-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:  57"
## [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
##   41   47   63   46   62   45   65   53   44   49   48   78  112   81  106
## 2011 2012
##  113  122
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   37   41   49   35   49   37   57   44   37   39   39   70   94   68   94
## 2011 2012
##   92  101
##

```

```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    37   40   49   35   48   33   56   42   37   39   39   65   86   66   86
## 2011 2012
##    86   96
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 19, df = 16, p-value = 0.3
```

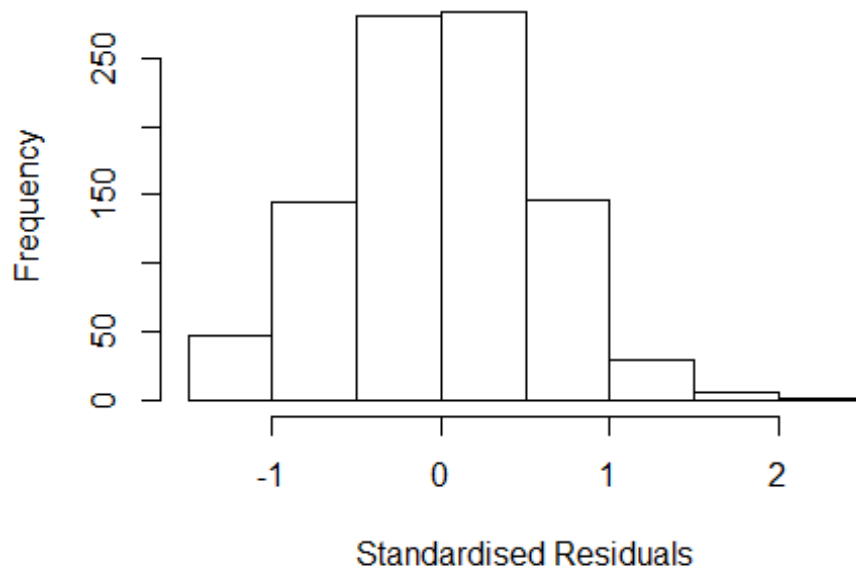


```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 5.1, df = 1, p-value = 0.02
```



```
## [1] "Female first author team size 2018 geometric mean: 1.62966142315199"
## [1] "Male first author team size 2018 geometric mean: 1.39725090154353"
##
## 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] "Female last author team size 2018 geometric mean: 1.61722184367085"
## [1] "Male last author team size 2018 geometric mean: 1.42045863373717"
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1800, 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 2.207 1      1.486
## LastAuthorFemale  2.203 1      1.484
## UniqueAuthors    1.557 4      1.057
## Year             1.639 16      1.016
```

Residuals from first and last author and team size



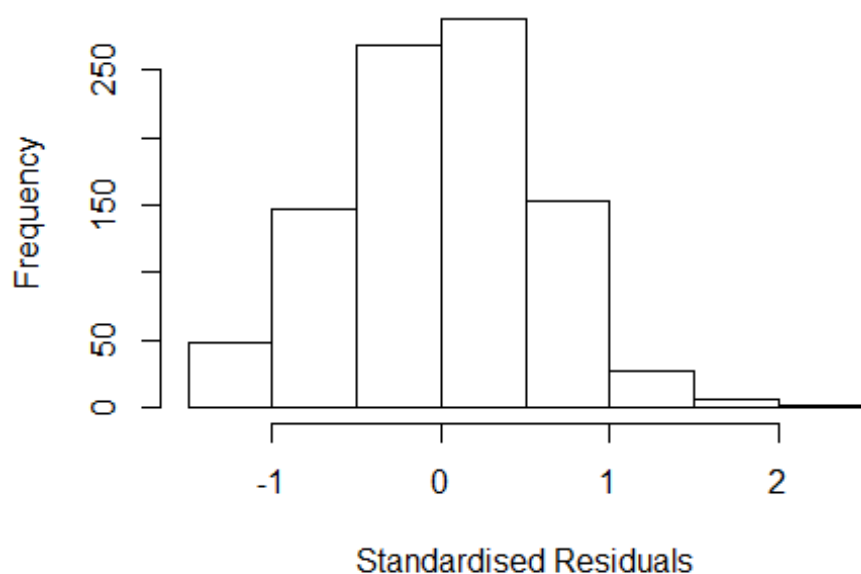
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.45988 -0.38490 -0.00627 0.40703 2.07228
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.817212 0.109662 7.45 2.1e-13 ***
## FirstAuthorFemale1 0.058494 0.056817 1.03 0.3035
## LastAuthorFemale1 0.000273 0.057576 0.00 0.9962
## UniqueAuthors2 0.131492 0.047803 2.75 0.0061 **
## UniqueAuthors3 0.362417 0.063223 5.73 1.3e-08 ***
## UniqueAuthors4 0.367491 0.131185 2.80 0.0052 **
## UniqueAuthors5 0.155455 0.131299 1.18 0.2367
## Year1997 0.153637 0.137246 1.12 0.2633
## Year1998 0.058031 0.142829 0.41 0.6846
## Year1999 0.049218 0.145347 0.34 0.7350
```

```

## Year2000          0.206911    0.151992    1.36    0.1737
## Year2001          0.056031    0.149481    0.37    0.7079
## Year2002          0.224068    0.128577    1.74    0.0817 .
## Year2003          0.121534    0.154378    0.79    0.4313
## Year2004         -0.035557    0.156965   -0.23    0.8208
## Year2005          0.059419    0.134535    0.44    0.6588
## Year2006          0.310508    0.135247    2.30    0.0219 *
## Year2007          0.226657    0.124278    1.82    0.0685 .
## Year2008          0.120688    0.123737    0.98    0.3296
## Year2009          0.221752    0.139014    1.60    0.1110
## Year2010          0.218896    0.120198    1.82    0.0689 .
## Year2011          0.301254    0.126705    2.38    0.0176 *
## Year2012          0.301451    0.124534    2.42    0.0157 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.595
## Multiple R-squared:  0.0734, Adjusted R-squared:  0.0511
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 69 weights are ~= 1. The remaining 871 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.201  0.873  0.953  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          1.06e-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 2.141 1          1.463
## LastAuthorFemale  2.254 1          1.501
## Year              1.152 16          1.004

```


Residuals from first and last author



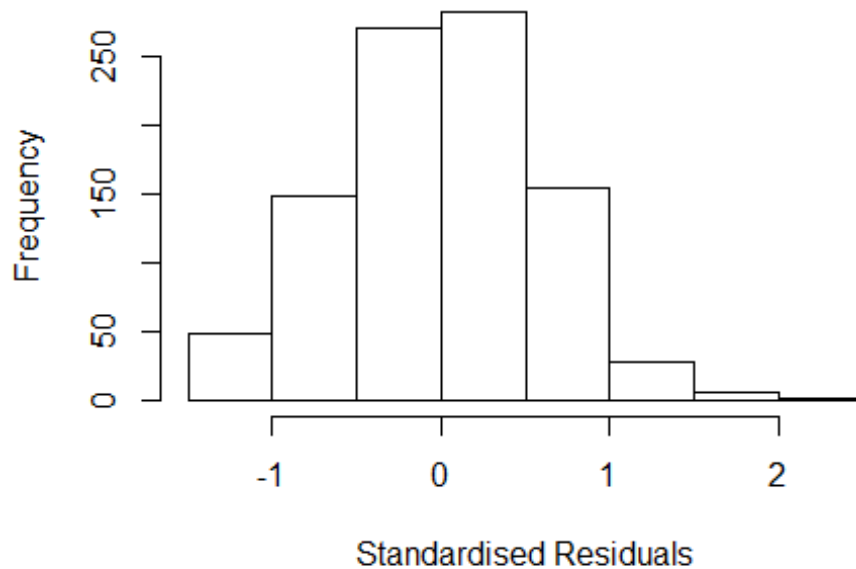
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.31321 -0.37949 0.00799 0.41306 2.33180
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8508 0.1136 7.49 1.6e-13 ***
## FirstAuthorFemale1 0.1172 0.0577 2.03 0.0425 *
## LastAuthorFemale1 -0.0381 0.0594 -0.64 0.5210
## Year1997 0.1424 0.1421 1.00 0.3163
## Year1998 0.0532 0.1427 0.37 0.7094
## Year1999 0.0380 0.1504 0.25 0.8005
## Year2000 0.2126 0.1582 1.34 0.1794
## Year2001 0.0816 0.1524 0.54 0.5927
## Year2002 0.2584 0.1304 1.98 0.0478 *
## Year2003 0.1237 0.1557 0.79 0.4269
## Year2004 -0.0218 0.1609 -0.14 0.8925
## Year2005 0.0853 0.1424 0.60 0.5494
```

```

## Year2006          0.3421      0.1400      2.44      0.0147 *
## Year2007          0.2572      0.1291      1.99      0.0467 *
## Year2008          0.1441      0.1277      1.13      0.2594
## Year2009          0.2691      0.1418      1.90      0.0580 .
## Year2010          0.2508      0.1256      2.00      0.0462 *
## Year2011          0.3215      0.1322      2.43      0.0152 *
## Year2012          0.3452      0.1289      2.68      0.0075 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.611
## Multiple R-squared:  0.0414, Adjusted R-squared:  0.0227
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 60 weights are ~= 1. The remaining 880 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.113  0.881  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      1.06e-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 1.046 1      1.023
## Year              1.046 16      1.001

```

Residuals from first author



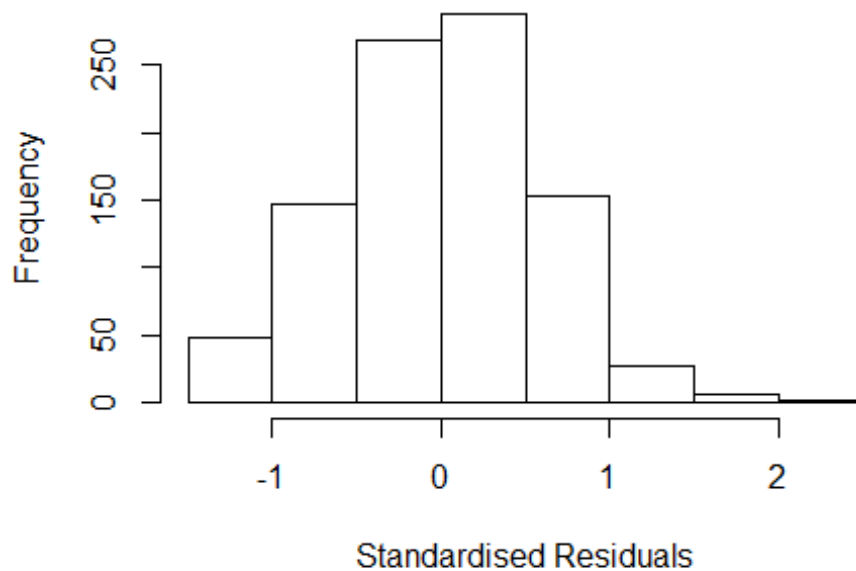
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.28288 -0.38100  0.00974  0.40686  2.29506
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.8503    0.1142    7.44 2.2e-13 ***
## FirstAuthorFemale1 0.0898    0.0404    2.22  0.0264 *
## Year1997        0.1396    0.1425    0.98  0.3275
## Year1998        0.0503    0.1433    0.35  0.7254
## Year1999        0.0364    0.1517    0.24  0.8106
## Year2000        0.2104    0.1586    1.33  0.1850
## Year2001        0.0800    0.1524    0.53  0.5995
## Year2002        0.2588    0.1310    1.98  0.0485 *
## Year2003        0.1209    0.1563    0.77  0.4393
## Year2004       -0.0230    0.1616   -0.14  0.8871
## Year2005        0.0824    0.1432    0.58  0.5654
## Year2006        0.3355    0.1395    2.41  0.0163 *
```

```

## Year2007          0.2521      0.1293      1.95      0.0515 .
## Year2008          0.1389      0.1281      1.08      0.2783
## Year2009          0.2643      0.1421      1.86      0.0632 .
## Year2010          0.2474      0.1262      1.96      0.0501 .
## Year2011          0.3207      0.1330      2.41      0.0161 *
## Year2012          0.3428      0.1296      2.65      0.0083 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.61
## Multiple R-squared:  0.0411, Adjusted R-squared:  0.0235
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 66 weights are ~= 1. The remaining 874 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.126  0.879   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      1.06e-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 1.097 1      1.047
## Year              1.097 16      1.003

```

Residuals from last author



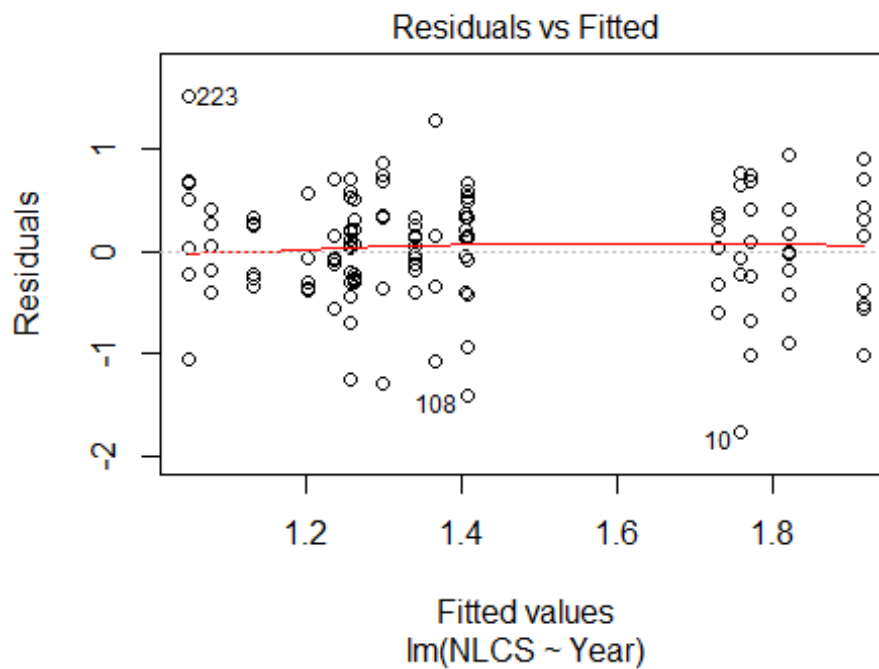
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.26307 -0.38436 0.00105 0.41658 2.22271
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8743 0.1145 7.64 5.6e-14 ***
## LastAuthorFemale1 0.0490 0.0415 1.18 0.2382
## Year1997 0.1323 0.1442 0.92 0.3594
## Year1998 0.0435 0.1442 0.30 0.7629
## Year1999 0.0262 0.1520 0.17 0.8634
## Year2000 0.2078 0.1600 1.30 0.1945
## Year2001 0.0715 0.1528 0.47 0.6400
## Year2002 0.2555 0.1320 1.94 0.0532 .
## Year2003 0.1094 0.1569 0.70 0.4858
## Year2004 -0.0281 0.1629 -0.17 0.8630
## Year2005 0.0792 0.1433 0.55 0.5805
## Year2006 0.3362 0.1407 2.39 0.0170 *
```

```

## Year2007          0.2487      0.1307      1.90      0.0574 .
## Year2008          0.1337      0.1293      1.03      0.3013
## Year2009          0.2571      0.1429      1.80      0.0723 .
## Year2010          0.2479      0.1270      1.95      0.0513 .
## Year2011          0.3200      0.1341      2.39      0.0172 *
## Year2012          0.3398      0.1302      2.61      0.0092 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.613
## Multiple R-squared:  0.0374, Adjusted R-squared:  0.0196
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 73 weights are ~= 1. The remaining 867 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.160  0.880  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      1.06e-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: 940"
## [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
##    8    8    9    7    6   10   11   10   10   15    8   11   12    7    9
## 2011 2012
##   21   13
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    7    6    6    4    6    8    9    8    6   11    8    9   11    6    7
## 2011 2012

```

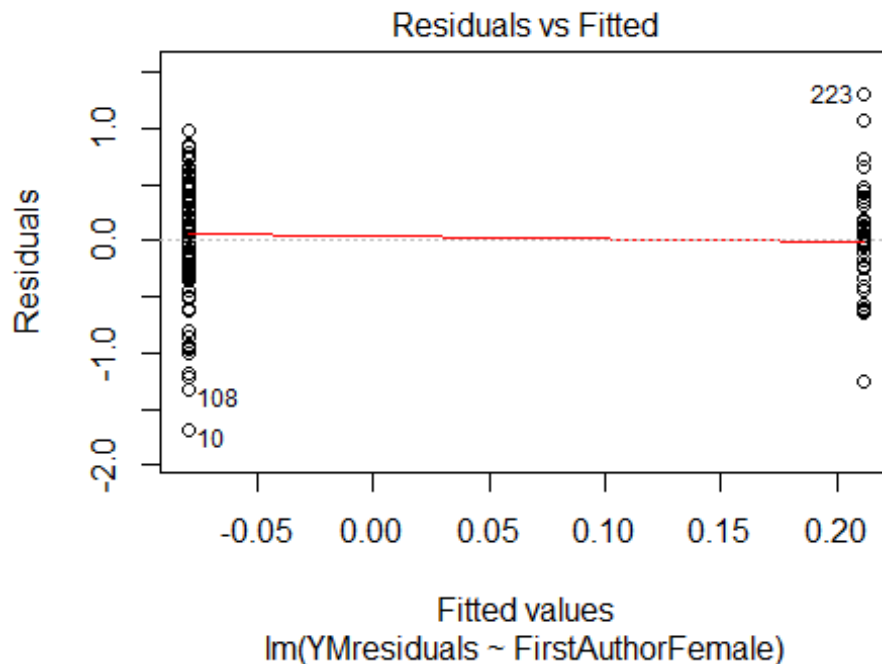
```
## 14 9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 7 5 5 4 6 6 8 7 6 10 8 7 10 6 6
## 2011 2012
## 12 9
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 42, df = 16, p-value = 5e-04
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.45, df = 1, p-value = 0.5
## [1] "Female first author team size 2018 geometric mean: 2.30505810033349"
## [1] "Male first author team size 2018 geometric mean: 3.55675950695046"
## 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 = 21, p-value = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.09127910518255"
## [1] "Male last author team size 2018 geometric mean: 3.57390741856396"

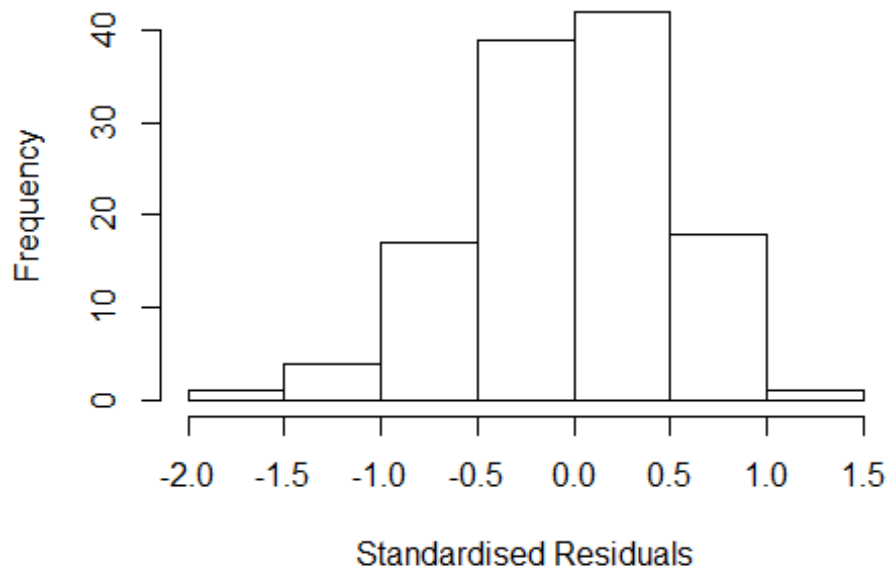
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 16, 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	3.620	1	1.903
LastAuthorFemale	2.163	1	1.471
UniqueAuthors	24.576	4	1.492
Year	43.309	16	1.125

Residuals from first and last author and team size



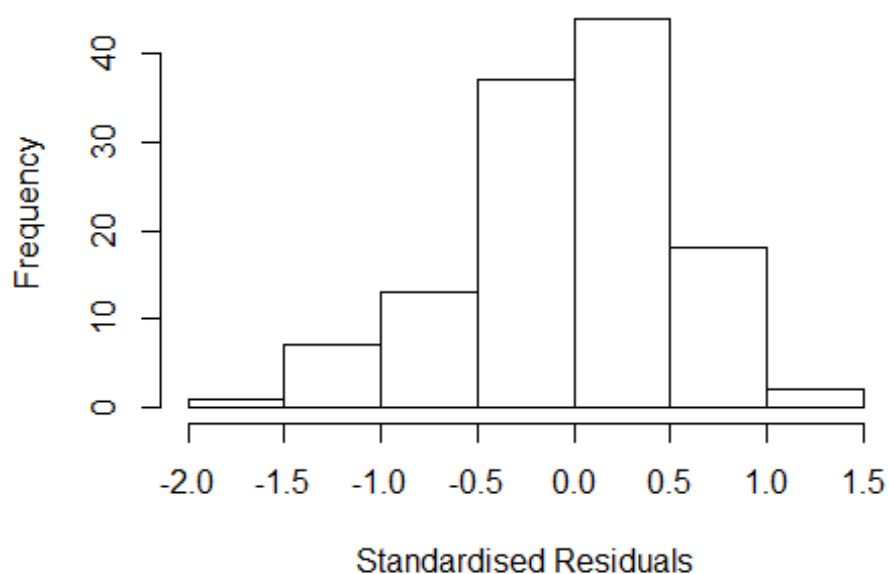
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.81394 -0.31835 0.00267 0.34313 1.23839
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7472 0.2945 5.93 4.4e-08 ***
## FirstAuthorFemale1 0.2960 0.1531 1.93 0.056 .
## LastAuthorFemale1 -0.0339 0.1459 -0.23 0.817
## UniqueAuthors2 0.1285 0.1788 0.72 0.474
## UniqueAuthors3 -0.1508 0.1509 -1.00 0.320
## UniqueAuthors4 0.5500 0.2457 2.24 0.027 *
## UniqueAuthors5 0.2760 0.2541 1.09 0.280
## Year1997 0.0668 0.5543 0.12 0.904
## Year1998 -0.0869 0.3360 -0.26 0.796
## Year1999 -0.6036 0.6618 -0.91 0.364
```

```

## Year2000          -0.5387      0.3440    -1.57      0.120
## Year2001          -0.3004      0.3804    -0.79      0.432
## Year2002           0.0414      0.3587     0.12      0.908
## Year2003          -0.4256      0.6422    -0.66      0.509
## Year2004          -0.6627      0.3223    -2.06      0.042 *
## Year2005          -0.3203      0.3544    -0.90      0.368
## Year2006          -0.6224      0.3019    -2.06      0.042 *
## Year2007          -0.4524      0.3172    -1.43      0.157
## Year2008          -0.6220      0.2977    -2.09      0.039 *
## Year2009          -0.6693      0.3001    -2.23      0.028 *
## Year2010          -0.7402      0.3459    -2.14      0.035 *
## Year2011          -0.6922      0.3236    -2.14      0.035 *
## Year2012          -0.9180      0.3895    -2.36      0.020 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.545
## Multiple R-squared:  0.248, Adjusted R-squared:  0.0813
## Convergence in 22 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 114 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.246  0.894  0.963   0.913   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      8.20e-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 2.329  1          1.526
## LastAuthorFemale  1.865  1          1.366
## Year              3.952 16          1.044

```

Residuals from first and last author



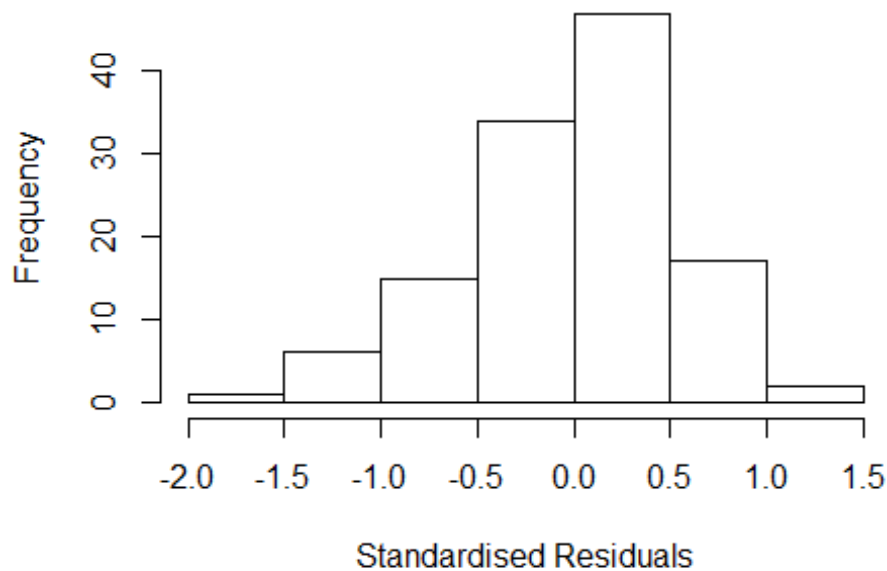
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.8012 -0.3020 0.0162 0.3292 1.3297
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7848 0.2835 6.29 7.6e-09 ***
## FirstAuthorFemale1 0.2757 0.1407 1.96 0.053 .
## LastAuthorFemale1 0.0626 0.1430 0.44 0.662
## Year1997 0.0164 0.5569 0.03 0.977
## Year1998 -0.1244 0.3254 -0.38 0.703
## Year1999 -0.5142 0.5608 -0.92 0.361
## Year2000 -0.6193 0.3332 -1.86 0.066 .
## Year2001 -0.2226 0.3746 -0.59 0.554
## Year2002 -0.0191 0.3644 -0.05 0.958
## Year2003 -0.5410 0.5372 -1.01 0.316
## Year2004 -0.7123 0.3321 -2.14 0.034 *
## Year2005 -0.3589 0.3629 -0.99 0.325
```

```

## Year2006          -0.5785      0.2958    -1.96     0.053 .
## Year2007          -0.4813      0.3071    -1.57     0.120
## Year2008          -0.5945      0.2945    -2.02     0.046 *
## Year2009          -0.7166      0.3089    -2.32     0.022 *
## Year2010          -0.8224      0.3259    -2.52     0.013 *
## Year2011          -0.6456      0.3265    -1.98     0.051 .
## Year2012          -0.8303      0.4950    -1.68     0.096 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.538
## Multiple R-squared:  0.207, Adjusted R-squared:  0.0685
## Convergence in 20 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 110 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.239 0.887 0.954 0.903 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      8.20e-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 2.059 1      1.435
## Year              2.059 16      1.023

```

Residuals from first author



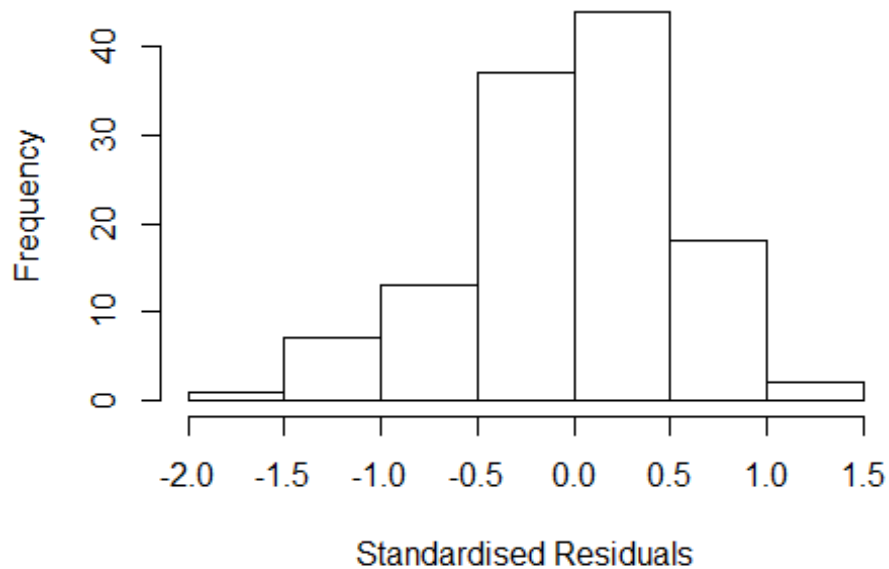
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.8002 -0.3014 0.0227 0.3576 1.3158
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.80281 0.28849 6.25 9.2e-09 ***
## FirstAuthorFemale1 0.29758 0.13560 2.19 0.030 *
## Year1997 -0.00263 0.51518 -0.01 0.996
## Year1998 -0.14260 0.32947 -0.43 0.666
## Year1999 -0.52915 0.56637 -0.93 0.352
## Year2000 -0.62978 0.33804 -1.86 0.065 .
## Year2001 -0.21185 0.37213 -0.57 0.570
## Year2002 -0.04051 0.36325 -0.11 0.911
## Year2003 -0.56911 0.52936 -1.08 0.285
## Year2004 -0.70487 0.33985 -2.07 0.041 *
## Year2005 -0.37086 0.37169 -1.00 0.321
## Year2006 -0.58246 0.30521 -1.91 0.059 .
```

```

## Year2007          -0.49633    0.31347   -1.58    0.116
## Year2008          -0.61021    0.30203   -2.02    0.046 *
## Year2009          -0.71685    0.32284   -2.22    0.029 *
## Year2010          -0.83301    0.32657   -2.55    0.012 *
## Year2011          -0.65855    0.33257   -1.98    0.050 .
## Year2012          -0.85620    0.46693   -1.83    0.070 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.548
## Multiple R-squared:  0.205, Adjusted R-squared:  0.075
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 111 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.258  0.888  0.959  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      8.20e-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 2.169 1          1.473
## Year              2.169 16          1.024

```

Residuals from last author



```
## [1] "List of 0 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
## -1.8863 -0.3223 0.0338 0.3536 1.5424
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.7590 0.2750 6.40 4.6e-09 ***
## LastAuthorFemale1 0.1519 0.1470 1.03 0.304
## Year1997 0.1273 0.4709 0.27 0.787
## Year1998 -0.0982 0.3178 -0.31 0.758
## Year1999 -0.4966 0.6448 -0.77 0.443
## Year2000 -0.5616 0.3164 -1.78 0.079 .
## Year2001 -0.1366 0.3825 -0.36 0.722
## Year2002 0.0534 0.3581 0.15 0.882
## Year2003 -0.2964 0.5677 -0.52 0.603
## Year2004 -0.6465 0.3247 -1.99 0.049 *
## Year2005 -0.2794 0.3364 -0.83 0.408
## Year2006 -0.5403 0.2869 -1.88 0.062 .
```

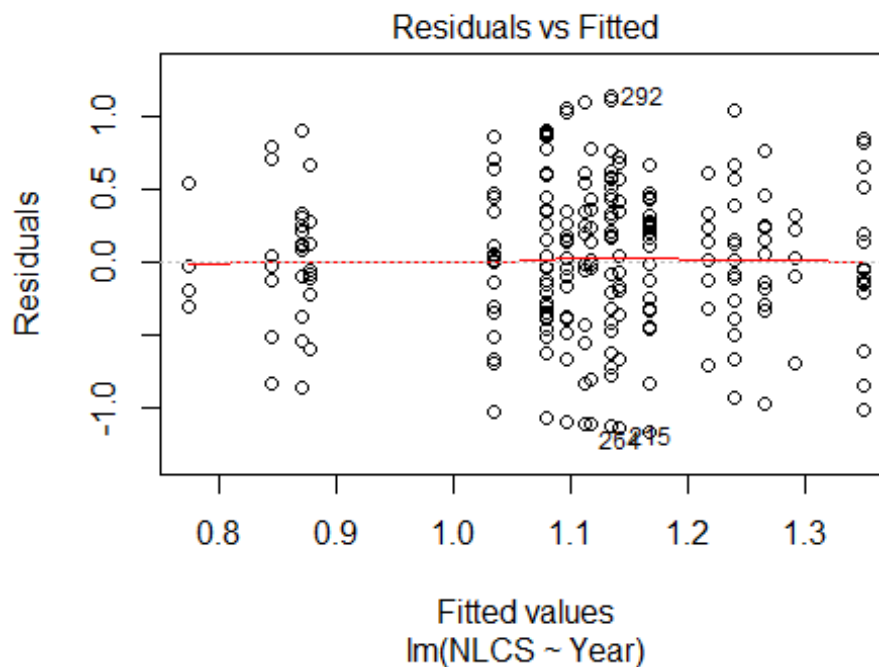
```

## Year2007          -0.3846      0.2997   -1.28    0.202
## Year2008          -0.4510      0.2766   -1.63    0.106
## Year2009          -0.6764      0.2847   -2.38    0.019 *
## Year2010          -0.6950      0.3025   -2.30    0.024 *
## Year2011          -0.5007      0.3032   -1.65    0.102
## Year2012          -0.7414      0.6354   -1.17    0.246
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.518
## Multiple R-squared:  0.191, Adjusted R-squared:  0.0584
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 114 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.156  0.901  0.953  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      8.20e-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: 122"
## [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
##   11   7  12  11  17  16  15  18   9  20  11  15  20  21  31
## 2011 2012
##   29  44
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   10   4   8   6  14  10  14  17   8  19  11  14  13  18  24
## 2011 2012

```



```
## 27 32
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 10 4 8 6 14 8 14 17 7 19 11 14 13 17 22
## 2011 2012
## 25 32
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 11, df = 16, p-value = 0.8
```



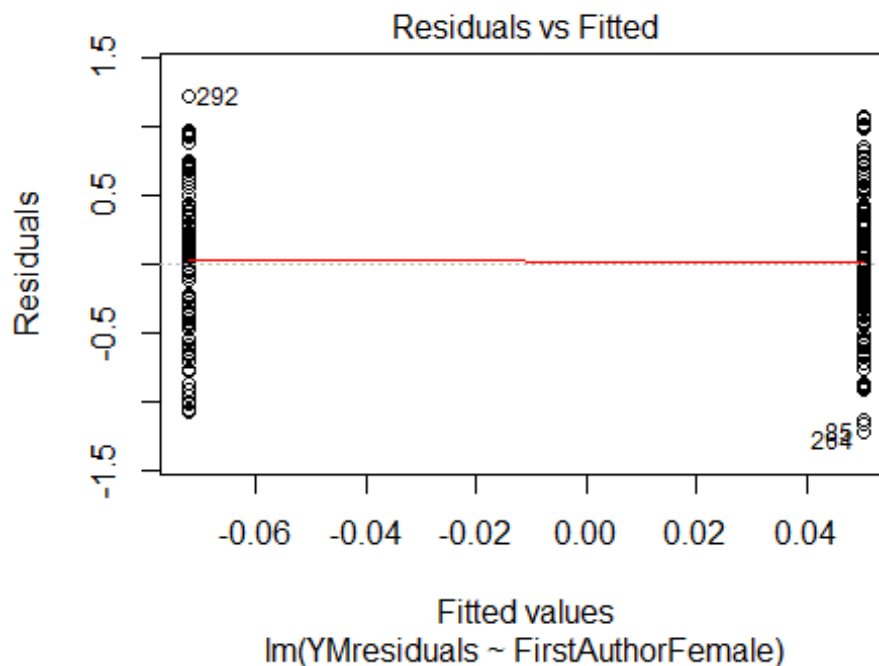
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 2.2, df = 1, p-value = 0.1

## [1] "Female first author team size 2018 geometric mean: 1.56653729645867"
## [1] "Male first author team size 2018 geometric mean: 1.44523016930896"

## 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 = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.57405761616385"
## [1] "Male last author team size 2018 geometric mean: 1.43356441463255"

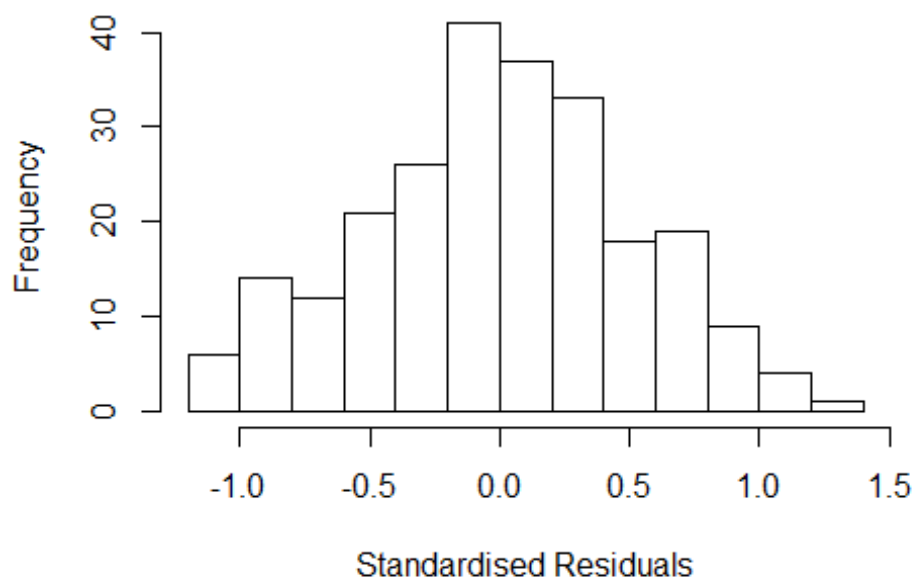
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 150, p-value = 0.7
## 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	2.835	1	1.684
LastAuthorFemale	3.382	1	1.839
UniqueAuthors	4.293	3	1.275
Year	5.721	16	1.056

Residuals from first and last author and team size



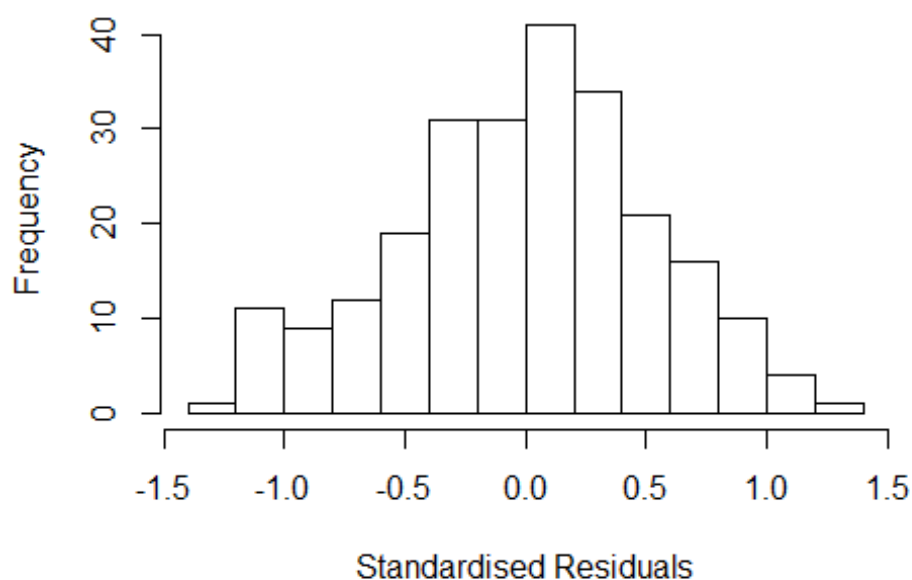
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1594 -0.3212 0.0058 0.2890 1.2863
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.09110 0.14843 7.35 3.9e-12 ***
## FirstAuthorFemale1 0.04348 0.11486 0.38 0.705
## LastAuthorFemale1 0.07672 0.11555 0.66 0.507
## UniqueAuthors2 0.08710 0.09016 0.97 0.335
## UniqueAuthors3 0.29124 0.13568 2.15 0.033 *
## UniqueAuthors4 0.28349 0.21540 1.32 0.190
## Year1997 -0.41214 0.21377 -1.93 0.055 .
## Year1998 -0.28490 0.23628 -1.21 0.229
## Year1999 0.14763 0.20072 0.74 0.463
## Year2000 -0.08522 0.20531 -0.42 0.679
```

```

## Year2001      -0.23319    0.27976   -0.83    0.405
## Year2002      0.01991    0.20513    0.10    0.923
## Year2003     -0.09523    0.18086   -0.53    0.599
## Year2004     -0.39732    0.18021   -2.20    0.029 *
## Year2005      0.15052    0.18238    0.83    0.410
## Year2006      0.00983    0.19614    0.05    0.960
## Year2007     -0.29127    0.19277   -1.51    0.132
## Year2008     -0.07178    0.20850   -0.34    0.731
## Year2009     -0.19016    0.18575   -1.02    0.307
## Year2010     -0.05190    0.16724   -0.31    0.757
## Year2011     -0.10242    0.20412   -0.50    0.616
## Year2012     -0.09990    0.18382   -0.54    0.587
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.0983, Adjusted R-squared:  0.0118
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 226 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.536  0.873   0.963   0.914   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      4.15e-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 3.116 1 1.765
## LastAuthorFemale 3.760 1 1.939
## Year 1.905 16 1.020

```

Residuals from first and last author



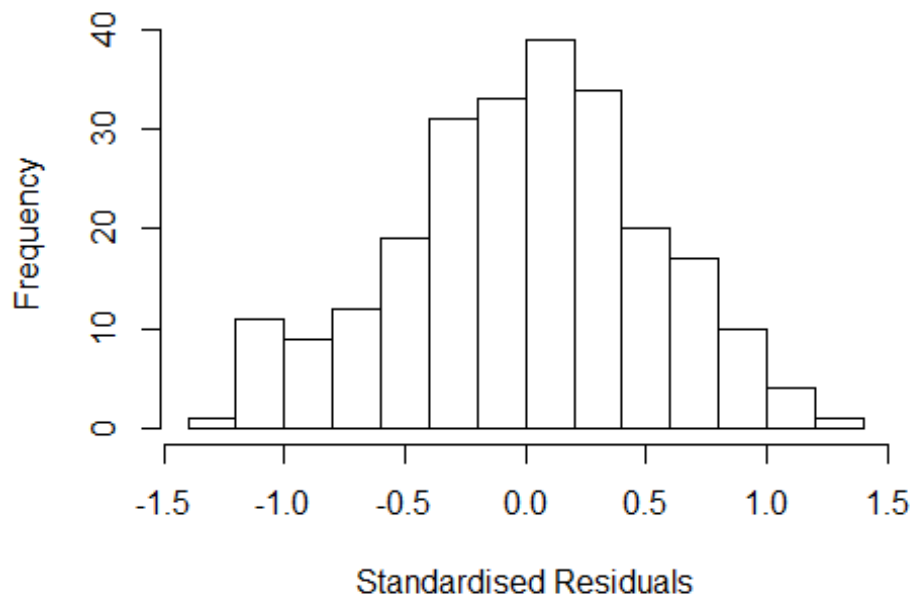
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2138 -0.3349 0.0445 0.3128 1.2589
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1408 0.1351 8.44 4e-15 ***
## FirstAuthorFemale1 0.0654 0.1228 0.53 0.595
## LastAuthorFemale1 0.0585 0.1255 0.47 0.642
## Year1997 -0.4220 0.1984 -2.13 0.034 *
## Year1998 -0.3355 0.2275 -1.47 0.142
## Year1999 0.1088 0.1917 0.57 0.571
## Year2000 -0.0949 0.2032 -0.47 0.641
## Year2001 -0.1325 0.2530 -0.52 0.601
## Year2002 0.0126 0.2035 0.06 0.951
## Year2003 -0.1264 0.1746 -0.72 0.470
## Year2004 -0.4263 0.1645 -2.59 0.010 *
## Year2005 0.1292 0.1773 0.73 0.467
```

```

## Year2006          0.0434      0.1757      0.25      0.805
## Year2007         -0.3083      0.1772     -1.74      0.083 .
## Year2008         -0.0705      0.2011     -0.35      0.726
## Year2009         -0.2087      0.1773     -1.18      0.240
## Year2010         -0.0508      0.1575     -0.32      0.748
## Year2011         -0.1247      0.1923     -0.65      0.517
## Year2012         -0.1286      0.1765     -0.73      0.467
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.0786, Adjusted R-squared:  0.00388
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 227 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.551  0.875  0.958  0.912  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      4.15e-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 1.244 1      1.115
## Year              1.244 16      1.007

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2081 -0.3313  0.0376  0.3276  1.2484
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.15458    0.13072   8.83 3.1e-16 ***
## FirstAuthorFemale1 0.11170    0.07798   1.43  0.1534
## Year1997      -0.44576    0.19043  -2.34  0.0201 *
## Year1998      -0.34620    0.22689  -1.53  0.1285
## Year1999       0.09082    0.18577   0.49  0.6254
## Year2000      -0.10259    0.20215  -0.51  0.6123
## Year2001      -0.15723    0.23554  -0.67  0.5051
## Year2002      -0.00365    0.19666  -0.02  0.9852
## Year2003      -0.13332    0.17377  -0.77  0.4438
## Year2004      -0.43322    0.16273  -2.66  0.0083 **
## Year2005       0.12624    0.17871   0.71  0.4807
## Year2006       0.03887    0.17618   0.22  0.8256
```

```

## Year2007          -0.32604    0.17115   -1.90    0.0581 .
## Year2008          -0.07503    0.20089   -0.37    0.7092
## Year2009          -0.21993    0.17476   -1.26    0.2095
## Year2010          -0.05820    0.15616   -0.37    0.7097
## Year2011          -0.12795    0.19307   -0.66    0.5082
## Year2012          -0.13381    0.17542   -0.76    0.4464
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.529
## Multiple R-squared:  0.0777, Adjusted R-squared:  0.0074
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 228 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.557  0.873  0.957   0.913  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      4.15e-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 1.485 1          1.219
## Year              1.485 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 ~ 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 -0.3336  0.0451  0.2993  1.2542

```



```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1419    0.1340   8.52 2.4e-15 ***
## LastAuthorFemale1 0.1118    0.0793   1.41  0.160
## Year1997         -0.4030    0.1941  -2.08  0.039 *
## Year1998         -0.3335    0.2274  -1.47  0.144
## Year1999          0.1242    0.1855   0.67  0.504
## Year2000         -0.0870    0.1987  -0.44  0.662
## Year2001         -0.1077    0.2403  -0.45  0.654
## Year2002          0.0299    0.1958   0.15  0.879
## Year2003         -0.1207    0.1732  -0.70  0.487
## Year2004         -0.4206    0.1619  -2.60  0.010 *
## Year2005          0.1332    0.1747   0.76  0.447
## Year2006          0.0514    0.1739   0.30  0.768
## Year2007         -0.2959    0.1732  -1.71  0.089 .
## Year2008         -0.0625    0.1984  -0.32  0.753
## Year2009         -0.1991    0.1742  -1.14  0.255
## Year2010         -0.0450    0.1551  -0.29  0.772
## Year2011         -0.1211    0.1924  -0.63  0.530
## Year2012         -0.1236    0.1746  -0.71  0.480
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.53
## Multiple R-squared:  0.0776, Adjusted R-squared:  0.00731
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 226 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.555 0.876 0.961 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      4.15e-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: 241"
## [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
##    2    3    6    1   10   11    8    7   13   35   22   29   24   29   39
## 2011 2012
##   49   51
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    5    1    8    9    7    7   11   31   20   28   24   23   33
## 2011 2012
##   41   47
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    3    5    1    7    8    7    5   11   28   20   27   22   21   29
## 2011 2012
##   41   45
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.21063699758223"
## [1] "Male first author team size 2018 geometric mean: 1.29357188425995"

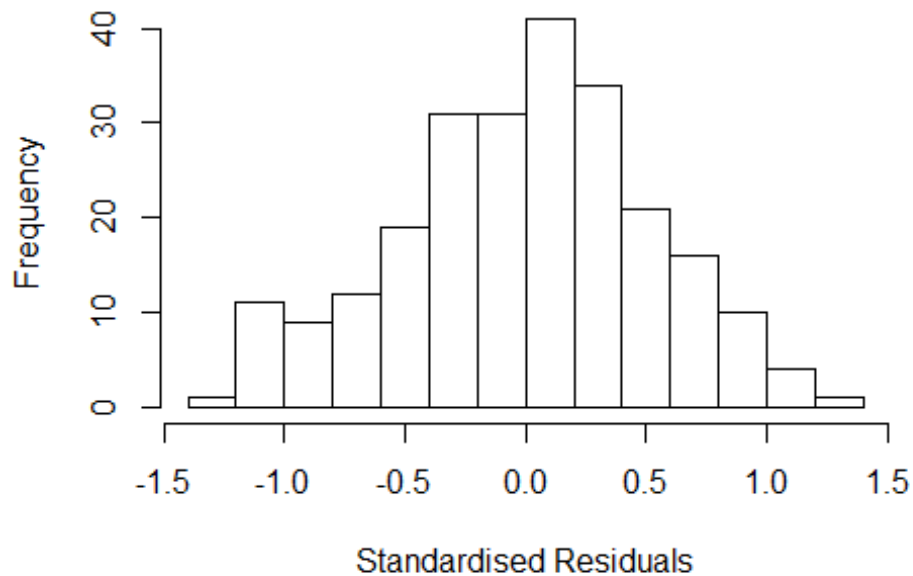
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.21063699758223"
## [1] "Male last author team size 2018 geometric mean: 1.29357188425995"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

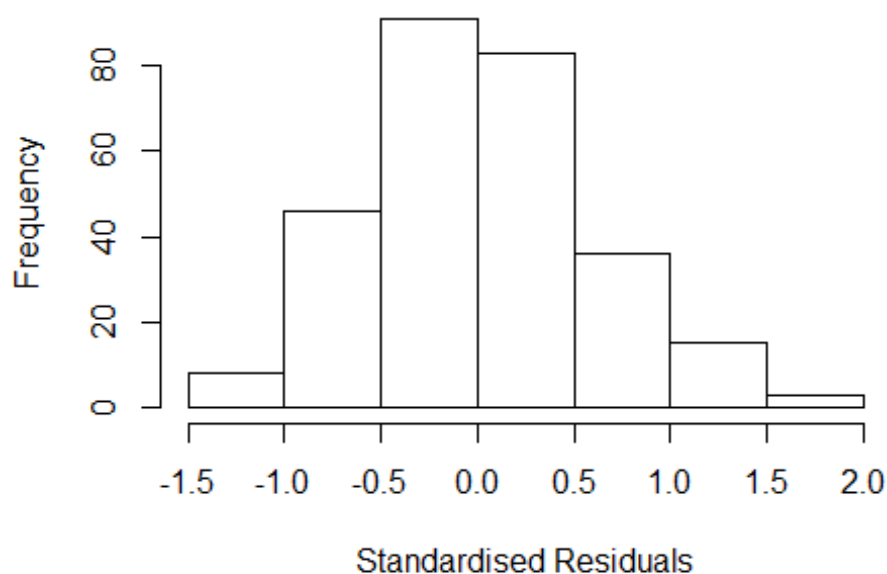
```

Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 130, 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 7.921 1          2.814
## LastAuthorFemale  7.590 1          2.755
## UniqueAuthors    4.705 4          1.214
## Year             11.839 16          1.080
```

Residuals from first and last author and team size



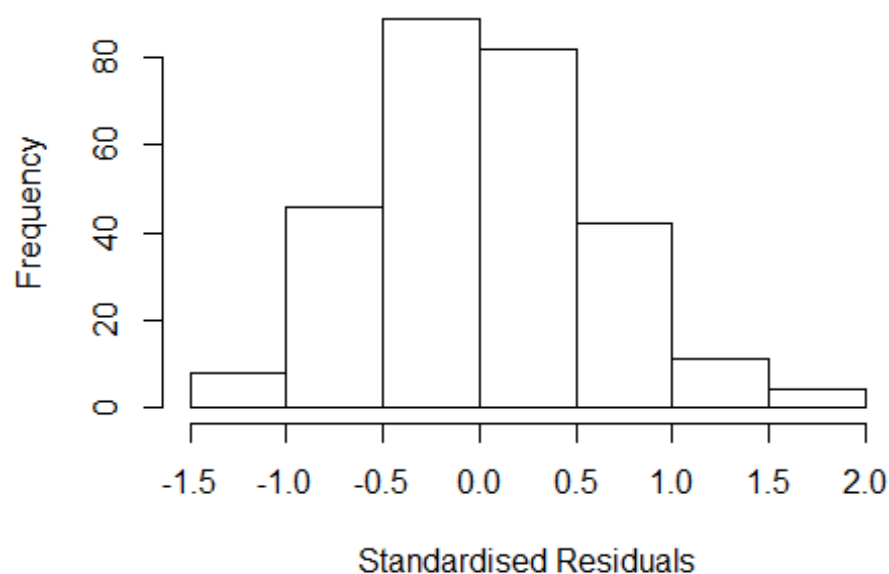
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.27662 -0.36689 -0.00855 0.35503 1.93775
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.88790 0.15243 5.82 1.7e-08 ***
## FirstAuthorFemale1 0.31071 0.11824 2.63 0.0091 **
## LastAuthorFemale1 -0.03811 0.12214 -0.31 0.7553
## UniqueAuthors2 0.17801 0.10730 1.66 0.0983 .
## UniqueAuthors3 0.17910 0.21696 0.83 0.4098
## UniqueAuthors4 0.44876 0.31223 1.44 0.1518
## UniqueAuthors5 0.12019 0.17959 0.67 0.5039
## Year1997 -0.49210 0.52912 -0.93 0.3532
## Year1998 -0.09133 0.25959 -0.35 0.7253
## Year1999 -0.57150 0.12405 -4.61 6.4e-06 ***
```

```

## Year2000      0.23263      0.25858      0.90      0.3691
## Year2001     -0.09853      0.26108     -0.38      0.7062
## Year2002      0.26216      0.17955      1.46      0.1455
## Year2003     -0.00761      0.27874     -0.03      0.9782
## Year2004      0.18570      0.23836      0.78      0.4366
## Year2005     -0.33035      0.16561     -1.99      0.0471 *
## Year2006     -0.08096      0.22878     -0.35      0.7237
## Year2007     -0.15465      0.19410     -0.80      0.4263
## Year2008     -0.06189      0.18709     -0.33      0.7410
## Year2009     -0.16365      0.17543     -0.93      0.3518
## Year2010     -0.12100      0.18737     -0.65      0.5190
## Year2011     -0.15718      0.17624     -0.89      0.3733
## Year2012     -0.10736      0.16473     -0.65      0.5152
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.571
## Multiple R-squared:  0.13,   Adjusted R-squared:  0.0558
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 251 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.225  0.851  0.951  0.903  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      3.55e-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 7.561 1      2.750
## LastAuthorFemale  6.323 1      2.515
## Year              4.216 16      1.046

```

Residuals from first and last author



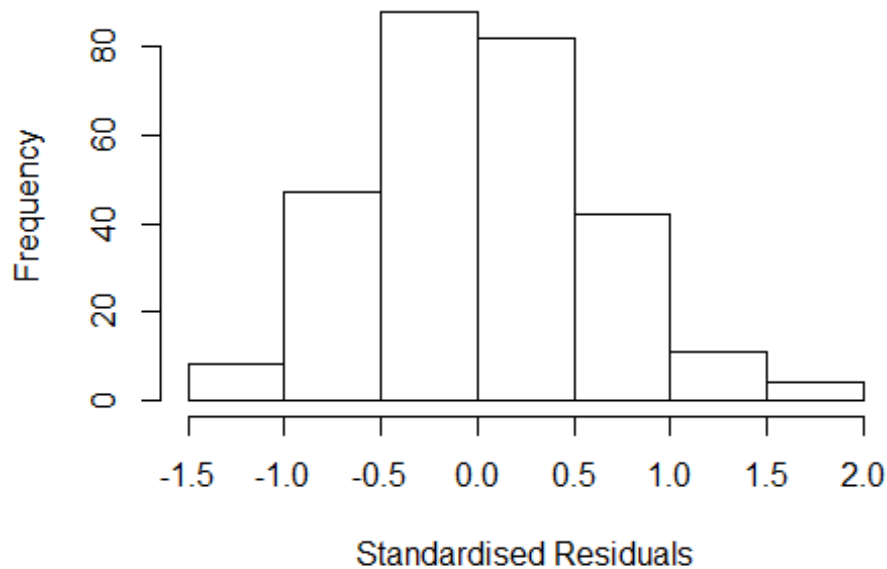
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.16499 -0.35427 -0.00606 0.35951 1.91477
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.85820 0.15086 5.69 3.4e-08 ***
## FirstAuthorFemale1 0.27243 0.11817 2.31 0.022 *
## LastAuthorFemale1 0.02987 0.11414 0.26 0.794
## Year1997 -0.45799 0.50941 -0.90 0.369
## Year1998 -0.08284 0.25809 -0.32 0.748
## Year1999 -0.57150 0.12393 -4.61 6.2e-06 ***
## Year2000 0.28390 0.24659 1.15 0.251
## Year2001 -0.03877 0.26217 -0.15 0.883
## Year2002 0.27906 0.18167 1.54 0.126
## Year2003 0.01986 0.27213 0.07 0.942
## Year2004 0.27270 0.22913 1.19 0.235
## Year2005 -0.27181 0.16422 -1.66 0.099 .
```

```

## Year2006      -0.01042    0.21479   -0.05    0.961
## Year2007      -0.06761    0.18079   -0.37    0.709
## Year2008       0.00449    0.18348    0.02    0.980
## Year2009      -0.11097    0.17861   -0.62    0.535
## Year2010      -0.02803    0.17703   -0.16    0.874
## Year2011      -0.11493    0.17490   -0.66    0.512
## Year2012      -0.05415    0.16499   -0.33    0.743
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.109, Adjusted R-squared:  0.0484
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 27 weights are ~= 1. The remaining 255 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.266  0.853   0.956   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      3.55e-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 3.561 1      1.887
## Year              3.561 16      1.040

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.16362 -0.35651 -0.00479 0.37025 1.90861
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.86550 0.14823 5.84 1.5e-08 ***
## FirstAuthorFemale1 0.29500 0.08130 3.63 0.00034 ***
## Year1997 -0.46555 0.51180 -0.91 0.36385
## Year1998 -0.08504 0.25852 -0.33 0.74244
## Year1999 -0.57150 0.12395 -4.61 6.3e-06 ***
## Year2000 0.28561 0.24845 1.15 0.25136
## Year2001 -0.04631 0.26079 -0.18 0.85918
## Year2002 0.27490 0.18028 1.52 0.12850
## Year2003 0.01327 0.27166 0.05 0.96108
## Year2004 0.26256 0.22243 1.18 0.23889
## Year2005 -0.27460 0.16364 -1.68 0.09451 .
## Year2006 -0.01131 0.21522 -0.05 0.95814
```

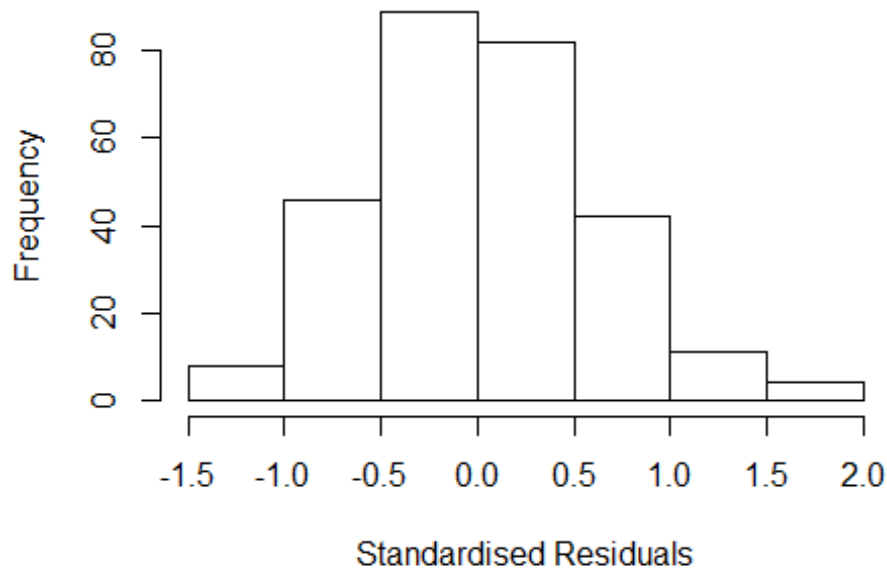


```

## Year2007          -0.07303    0.18072   -0.40  0.68648
## Year2008           0.00312    0.18350    0.02  0.98645
## Year2009          -0.11211    0.17852   -0.63  0.53053
## Year2010          -0.03307    0.17682   -0.19  0.85179
## Year2011          -0.11999    0.17411   -0.69  0.49134
## Year2012          -0.05810    0.16484   -0.35  0.72476
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.584
## Multiple R-squared:  0.109, Adjusted R-squared:  0.052
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 24 weights are ~= 1. The remaining 258 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.264  0.857  0.955  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      3.55e-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 2.887 1          1.699
## Year            2.887 16          1.034

```

Residuals from last author



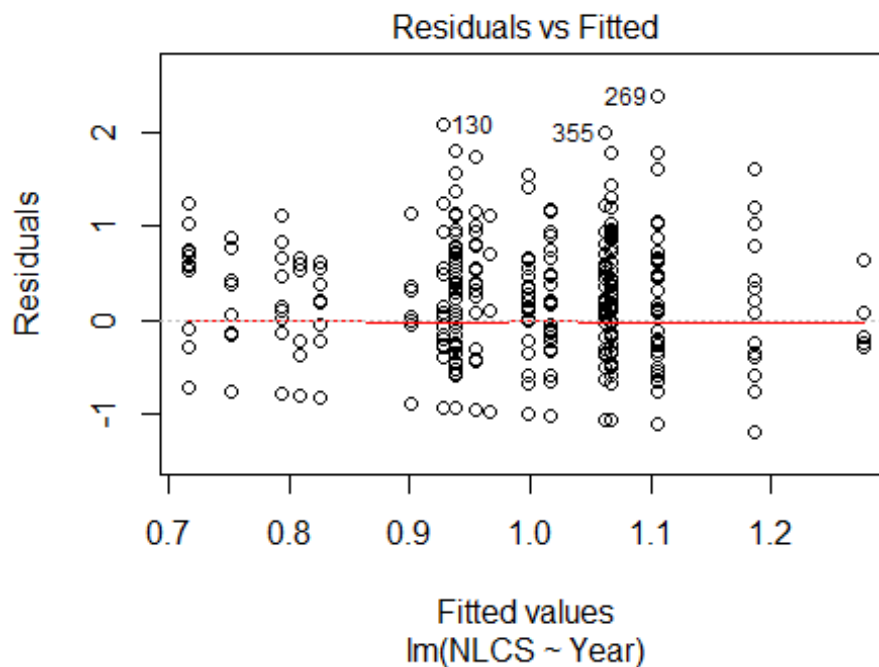
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.2127 -0.3844 -0.0016 0.3827 1.9242
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9267 0.1471 6.30 1.3e-09 ***
## LastAuthorFemale1 0.2338 0.0793 2.95 0.0035 **
## Year1997 -0.5211 0.5384 -0.97 0.3340
## Year1998 -0.1046 0.2600 -0.40 0.6878
## Year1999 -0.5715 0.1240 -4.61 6.3e-06 ***
## Year2000 0.2214 0.2381 0.93 0.3533
## Year2001 -0.1076 0.2601 -0.41 0.6795
## Year2002 0.2399 0.1747 1.37 0.1707
## Year2003 -0.0407 0.2801 -0.15 0.8845
## Year2004 0.2860 0.2304 1.24 0.2156
## Year2005 -0.3377 0.1641 -2.06 0.0406 *
## Year2006 -0.0523 0.2119 -0.25 0.8054
```

```

## Year2007          -0.0840      0.1807   -0.46   0.6425
## Year2008          -0.0657      0.1762   -0.37   0.7098
## Year2009          -0.1889      0.1709   -1.11   0.2699
## Year2010          -0.0702      0.1795   -0.39   0.6959
## Year2011          -0.1521      0.1721   -0.88   0.3776
## Year2012          -0.1072      0.1625   -0.66   0.5101
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.584
## Multiple R-squared:  0.0926, Adjusted R-squared:  0.0342
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 25 weights are ~= 1. The remaining 257 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.256  0.846  0.955  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      3.55e-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: 282"
## [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
##   11   12   14   11   11    7   17   18   24   26   19   34   54   40   51
## 2011 2012
##   80   79
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    9    9   11    7    5    5   12   18   21   24   18   27   45   33   42
## 2011 2012

```

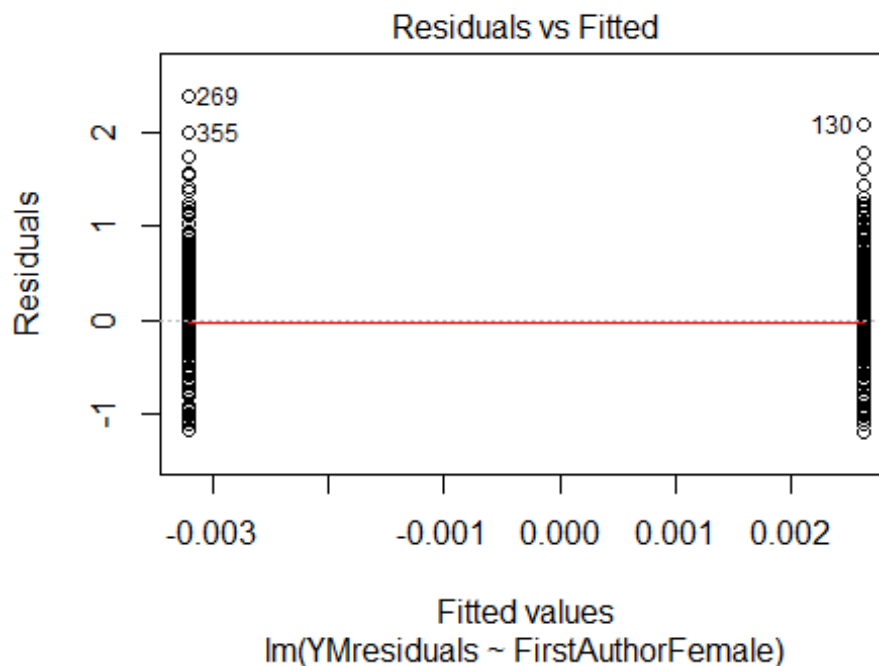
```
## 71 70
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 8 9 11 7 5 5 12 17 21 24 18 26 44 33 41
## 2011 2012
## 70 68
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 11, df = 16, p-value = 0.8
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.00097, df = 1, p-value = 1
## [1] "Female first author team size 2018 geometric mean: 1.3572711156556"
## [1] "Male first author team size 2018 geometric mean: 1.37359560712605"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

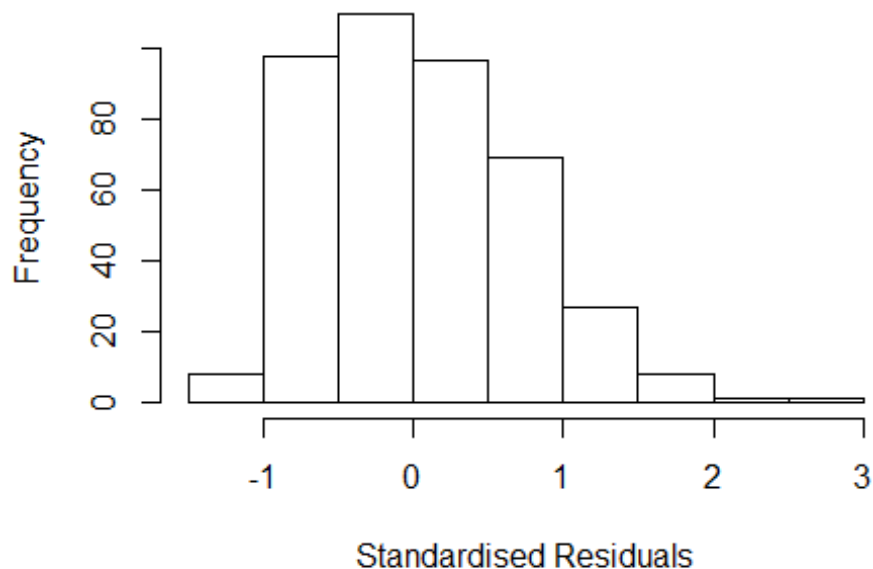
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 540, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.37387974235394"
## [1] "Male last author team size 2018 geometric mean: 1.34687176593718"

## 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 = 550, 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 4.171  1      2.042
## LastAuthorFemale  4.077  1      2.019
## UniqueAuthors    1.900  4      1.084
## Year              2.279 16      1.026
```

Residuals from first and last author and team size



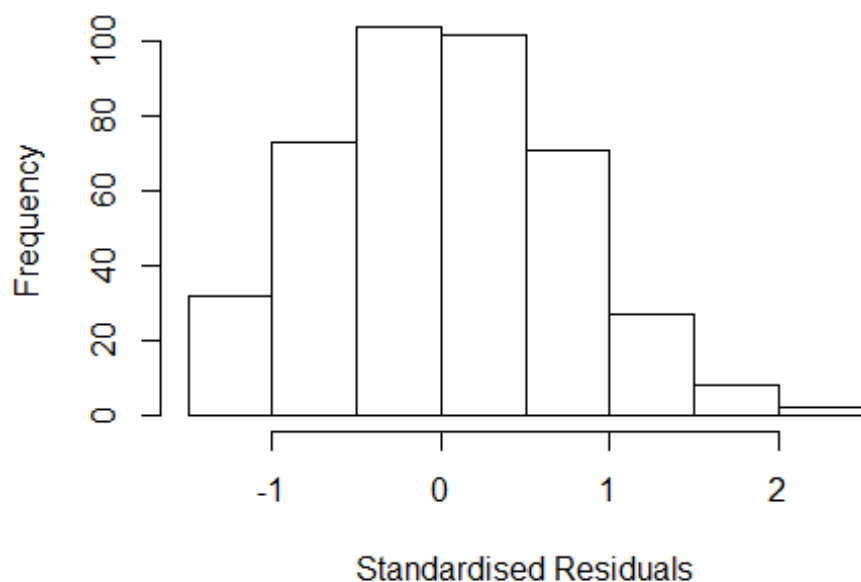
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 269 84998080688 3.482 2008      1213      2      2.524
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.3949 -0.5057 -0.0426  0.5068  2.5245
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.74273    0.22016   3.37 0.00082 ***
## FirstAuthorFemale1 0.08744    0.14007   0.62 0.53283
## LastAuthorFemale1 -0.09932    0.14054  -0.71 0.48017
## UniqueAuthors2    0.34254    0.12100   2.83 0.00488 **
## UniqueAuthors3   -0.12174    0.16706  -0.73 0.46659
## UniqueAuthors4    0.31544    0.30522   1.03 0.30200
## UniqueAuthors5    0.64600    0.11313   5.71 2.2e-08 ***
## Year1997          0.05440    0.27705   0.20 0.84443
## Year1998          0.02459    0.28522   0.09 0.93135
## Year1999          0.00534    0.28960   0.02 0.98529
```

```

## Year2000          0.53054      0.26759      1.98  0.04809 *
## Year2001          0.23143      0.53810      0.43  0.66737
## Year2002          0.05245      0.30407      0.17  0.86313
## Year2003         -0.05068      0.29898     -0.17  0.86549
## Year2004          0.05562      0.27622      0.20  0.84053
## Year2005          0.22063      0.26842      0.82  0.41159
## Year2006          0.44517      0.30365      1.47  0.14342
## Year2007          0.09077      0.26988      0.34  0.73681
## Year2008          0.22668      0.24599      0.92  0.35735
## Year2009          0.21111      0.23489      0.90  0.36932
## Year2010          0.24934      0.23706      1.05  0.29353
## Year2011          0.13468      0.23266      0.58  0.56301
## Year2012          0.24476      0.23978      1.02  0.30797
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.73
## Multiple R-squared:  0.0525, Adjusted R-squared:  -0.000143
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 32 weights are ~= 1. The remaining 387 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.208  0.873  0.945  0.915  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      2.39e-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 4.988 1      2.233
## LastAuthorFemale 4.988 1      2.233
## Year      1.227 16      1.006

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.17480 -0.51189 0.00381 0.55113 2.46765
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8419 0.2230 3.78 0.00018 ***
## FirstAuthorFemale1 0.0699 0.1596 0.44 0.66170
## LastAuthorFemale1 -0.0640 0.1598 -0.40 0.68911
## Year1997 -0.0144 0.2788 -0.05 0.95897
## Year1998 -0.0908 0.2846 -0.32 0.74998
## Year1999 -0.0453 0.3086 -0.15 0.88330
## Year2000 0.4246 0.2704 1.57 0.11715
## Year2001 0.1266 0.5300 0.24 0.81139
## Year2002 -0.0552 0.3032 -0.18 0.85572
## Year2003 -0.1406 0.3070 -0.46 0.64724
## Year2004 0.0127 0.2785 0.05 0.96358
## Year2005 0.1445 0.2693 0.54 0.59191
```

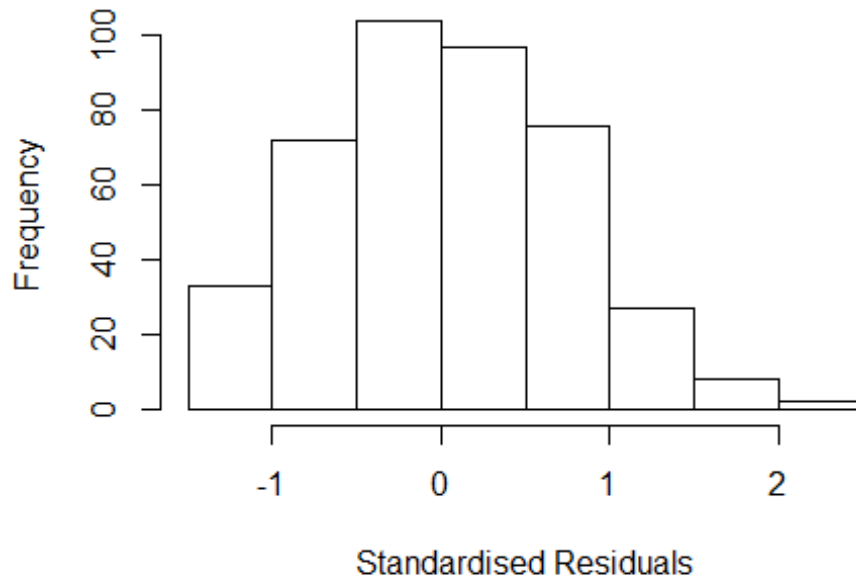


```

## Year2006          0.3270      0.3062      1.07  0.28613
## Year2007          0.0473      0.2800      0.17  0.86589
## Year2008          0.1666      0.2482      0.67  0.50260
## Year2009          0.1613      0.2456      0.66  0.51157
## Year2010          0.2070      0.2432      0.85  0.39517
## Year2011          0.0607      0.2383      0.25  0.79912
## Year2012          0.1834      0.2422      0.76  0.44937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.755
## Multiple R-squared:  0.0231, Adjusted R-squared:  -0.0209
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 21 weights are ~= 1. The remaining 398 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.263  0.876  0.950  0.920  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      2.39e-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 1.119 1      1.058
## Year              1.119 16      1.004

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.17326 -0.51209 0.00132 0.55609 2.46284
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.84690 0.22379 3.78 0.00018 ***
## FirstAuthorFemale1 0.01395 0.07558 0.18 0.85365
## Year1997 -0.01696 0.28111 -0.06 0.95192
## Year1998 -0.09720 0.28459 -0.34 0.73289
## Year1999 -0.04557 0.31370 -0.15 0.88457
## Year2000 0.41631 0.27097 1.54 0.12524
## Year2001 0.11907 0.53011 0.22 0.82240
## Year2002 -0.06417 0.30283 -0.21 0.83230
## Year2003 -0.14956 0.30638 -0.49 0.62571
## Year2004 0.00557 0.27880 0.02 0.98406
## Year2005 0.13276 0.26771 0.50 0.62023
## Year2006 0.31241 0.30290 1.03 0.30298
```

```

## Year2007          0.03947    0.28012    0.14  0.88801
## Year2008          0.15831    0.24771    0.64  0.52312
## Year2009          0.15624    0.24587    0.64  0.52549
## Year2010          0.20099    0.24275    0.83  0.40817
## Year2011          0.05022    0.23709    0.21  0.83237
## Year2012          0.17214    0.24214    0.71  0.47756
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.756
## Multiple R-squared:  0.0226, Adjusted R-squared:  -0.0189
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 23 weights are ~= 1. The remaining 396 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.267  0.876  0.950  0.920  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      2.39e-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 1.12  1      1.058
## Year              1.12 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
## -1.17013 -0.51573 -0.00556  0.55762  2.47170

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.85491    0.22379   3.82 0.00015 ***
## LastAuthorFemale1 -0.00253    0.07564  -0.03 0.97333
## Year1997        -0.01597    0.28411  -0.06 0.95519
## Year1998        -0.10231    0.28585  -0.36 0.72060
## Year1999        -0.04458    0.31339  -0.14 0.88695
## Year2000         0.41518    0.27134   1.53 0.12678
## Year2001         0.11602    0.52776   0.22 0.82611
## Year2002        -0.06390    0.30461  -0.21 0.83394
## Year2003        -0.14896    0.30822  -0.48 0.62914
## Year2004         0.00370    0.28004   0.01 0.98945
## Year2005         0.13238    0.26897   0.49 0.62286
## Year2006         0.31522    0.30651   1.03 0.30438
## Year2007         0.03797    0.28109   0.14 0.89261
## Year2008         0.15792    0.24936   0.63 0.52689
## Year2009         0.15480    0.24760   0.63 0.53220
## Year2010         0.20151    0.24518   0.82 0.41163
## Year2011         0.04964    0.23924   0.21 0.83572
## Year2012         0.17342    0.24483   0.71 0.47915
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.763
## Multiple R-squared:  0.0224, Adjusted R-squared:  -0.019
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 22 weights are ~= 1. The remaining 397 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.272 0.877  0.951  0.922  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      2.39e-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: 419"
## [1] ""

```

```

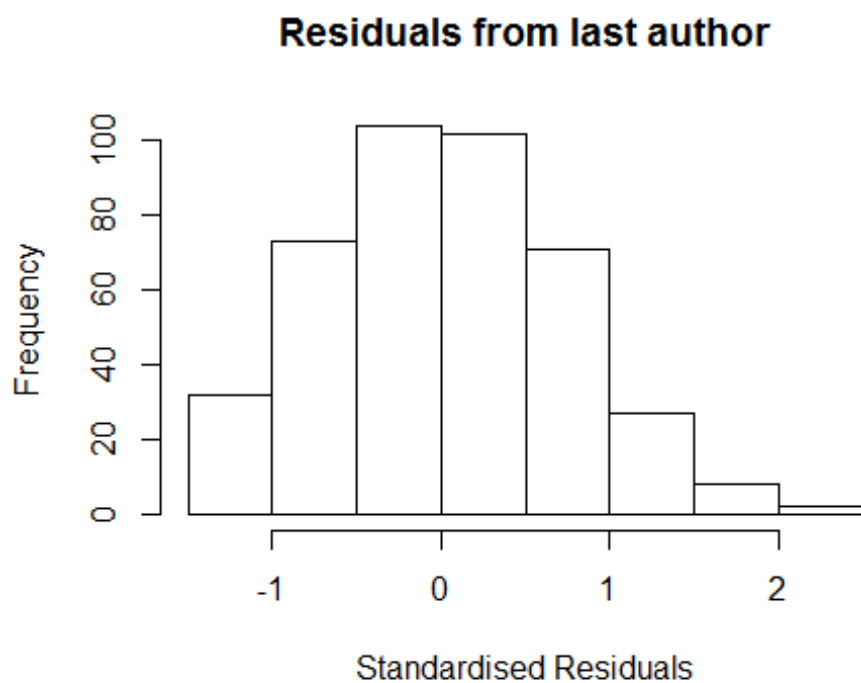
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3317"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    2    2    6    7    4    2    7    4    8    6    5    8    9    3
## 2011 2012
##    9    14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    2    1    4    2    2    6    3    4    5    5    7    7    3
## 2011 2012
##    7    11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    2    1    4    2    2    5    3    4    5    4    7    7    3
## 2011 2012
##    7    11
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1"
## [1] "Male first author team size 2018 geometric mean: 1.14869835499704"

## 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 = 0.5
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.14869835499704"
## [1] "Male last author team size 2018 geometric mean: 1"

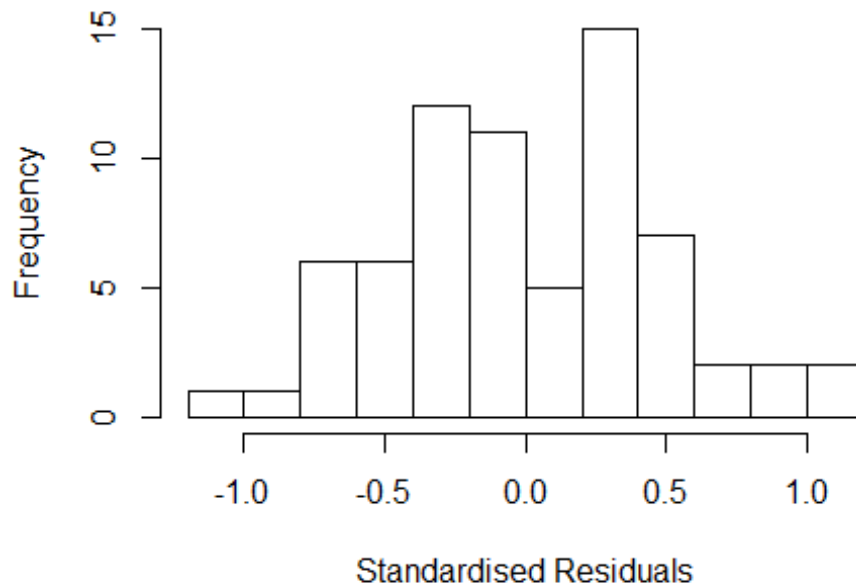
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, 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 9.157e+13 1      9.569e+06
## LastAuthorFemale  1.105e+02 1      1.051e+01
## UniqueAuthors    -2.122e+28 3           NaN
## Year              -5.120e+30 16           NaN
```

Residuals from first and last author and team size



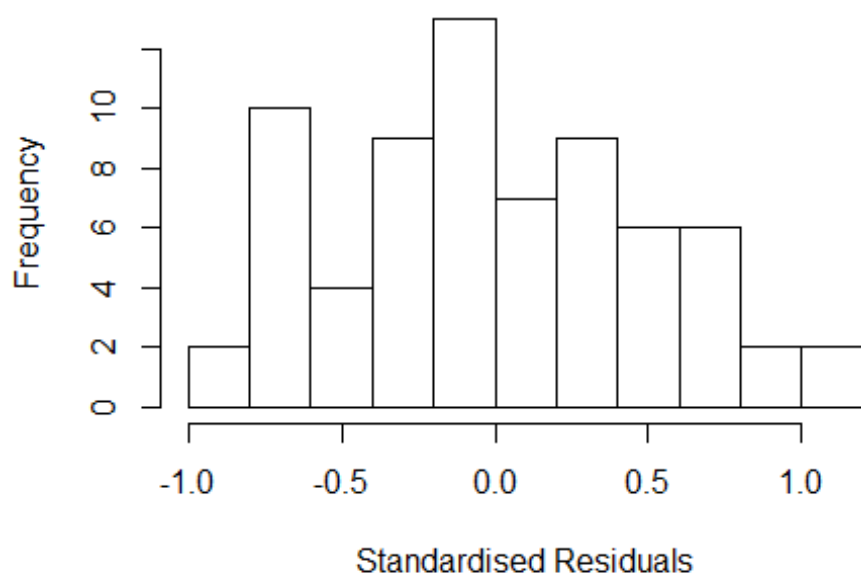
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.03e+00 -3.18e-01 5.55e-17 3.21e-01 1.09e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.91739 0.54737 1.68 0.1002
## FirstAuthorFemale1 0.74631 0.33945 2.20 0.0328 *
## LastAuthorFemale1 -0.24109 0.27865 -0.87 0.3912
## UniqueAuthors2 0.33376 0.15929 2.10 0.0414 *
## UniqueAuthors3 0.71354 0.49580 1.44 0.1566
## UniqueAuthors4 -0.62053 0.19670 -3.15 0.0028 **
## Year1997 -0.91739 0.54737 -1.68 0.1002
## Year1998 0.16061 0.77016 0.21 0.8357
## Year1999 -0.66346 0.59628 -1.11 0.2714
## Year2000 -0.07961 0.69744 -0.11 0.9096
```

```

## Year2001      -0.47000      0.70428      -0.67      0.5077
## Year2002      0.43650      0.54088      0.81      0.4236
## Year2003     -0.53360      0.59937     -0.89      0.3778
## Year2004     -0.41472      0.58938     -0.70      0.4850
## Year2005     -0.05652      0.56859     -0.10      0.9212
## Year2006     -0.42860      0.58481     -0.73      0.4672
## Year2007     -0.61881      0.55960     -1.11      0.2743
## Year2008     -0.09803      0.56864     -0.17      0.8638
## Year2009     -0.75461      0.57317     -1.32      0.1942
## Year2010     -0.36353      0.66590     -0.55      0.5877
## Year2011     -0.00516      0.58167     -0.01      0.9930
## Year2012     -0.29686      0.57323     -0.52      0.6069
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.615
## Multiple R-squared:  0.4, Adjusted R-squared:  0.138
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 58 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.734  0.919  0.964  0.939  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      1.43e-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 3.843e+13 1      6.199e+06
## LastAuthorFemale  4.715e+01 1      6.867e+00
## Year              7.304e+14 16      2.914e+00

```


Residuals from first and last author



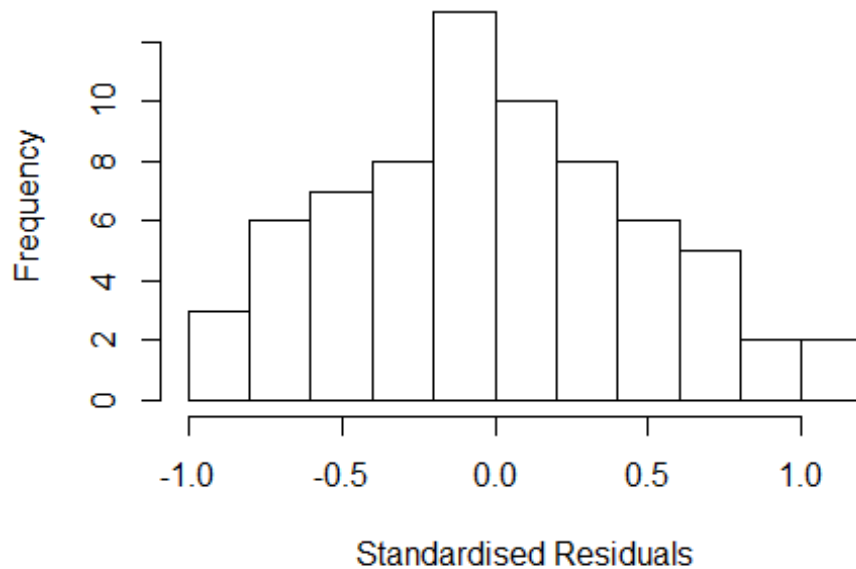
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.8274 -0.3308 -0.0241 0.3646 1.0906
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9058 0.5191 1.74 0.0871 .
## FirstAuthorFemale1 0.8589 0.2835 3.03 0.0038 **
## LastAuthorFemale1 -0.3305 0.2338 -1.41 0.1636
## Year1997 -0.9058 0.5191 -1.74 0.0871 .
## Year1998 0.1722 0.7388 0.23 0.8166
## Year1999 -0.4307 0.5667 -0.76 0.4508
## Year2000 -0.0784 0.6683 -0.12 0.9071
## Year2001 0.1100 0.6988 0.16 0.8756
## Year2002 0.4365 0.5119 0.85 0.3978
## Year2003 -0.4766 0.5848 -0.82 0.4188
## Year2004 -0.3323 0.5571 -0.60 0.5534
## Year2005 0.0132 0.5261 0.03 0.9800
```

```

## Year2006          -0.3762      0.5721   -0.66   0.5138
## Year2007          -0.6189      0.5315   -1.16   0.2497
## Year2008           0.0739      0.5451    0.14   0.8927
## Year2009          -0.6324      0.5529   -1.14   0.2581
## Year2010           0.1930      0.6303    0.31   0.7607
## Year2011           0.0582      0.5521    0.11   0.9164
## Year2012          -0.3216      0.5391   -0.60   0.5535
## ---
## 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:  0.112
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 60 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.763  0.917  0.957   0.943  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      1.43e-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 1.991e+14 1      1.411e+07
## Year              1.991e+14 16      2.798e+00

```

Residuals from first author



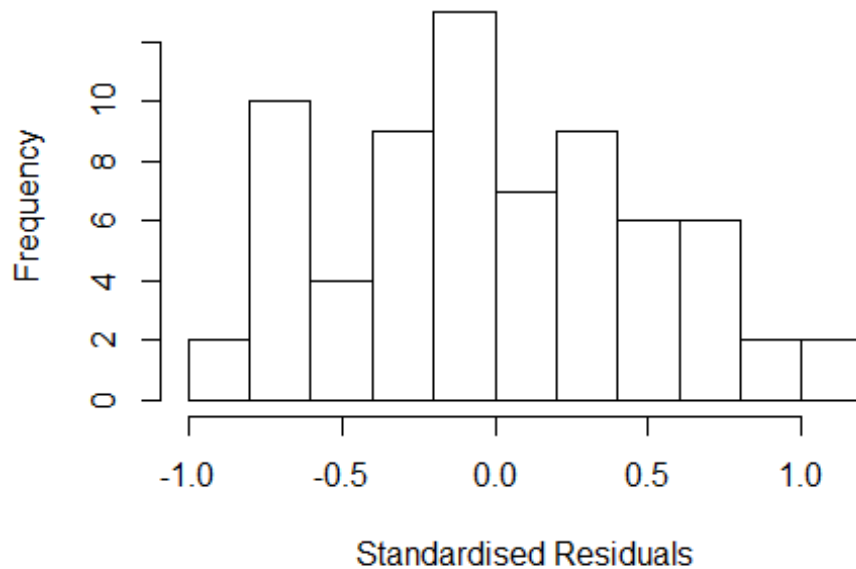
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.99137 -0.34207 -0.00223 0.36123 1.11574
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.88720 0.49836 1.78 0.0809 .
## FirstAuthorFemale1 0.56560 0.18118 3.12 0.0029 **
## Year1997 -0.88720 0.49836 -1.78 0.0809 .
## Year1998 0.19080 0.72454 0.26 0.7933
## Year1999 -0.11880 0.49794 -0.24 0.8124
## Year2000 -0.08494 0.64646 -0.13 0.8960
## Year2001 -0.03670 0.57945 -0.06 0.9497
## Year2002 0.43650 0.49098 0.89 0.3781
## Year2003 -0.49521 0.56729 -0.87 0.3867
## Year2004 -0.22149 0.53120 -0.42 0.6784
## Year2005 0.00391 0.50645 0.01 0.9939
## Year2006 -0.39473 0.55341 -0.71 0.4789
```

```

## Year2007      -0.61975    0.51112   -1.21    0.2308
## Year2008      0.11566    0.52113    0.22    0.8252
## Year2009     -0.59538    0.53785   -1.11    0.2734
## Year2010      0.10417    0.70934    0.15    0.8838
## Year2011      0.06419    0.52968    0.12    0.9040
## Year2012     -0.31964    0.51797   -0.62    0.5399
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.654
## Multiple R-squared:  0.331, Adjusted R-squared:  0.112
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 61 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.753  0.918  0.958  0.943  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      1.43e-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 12.3 1          3.507
## Year              12.3 16          1.082

```

Residuals from last author



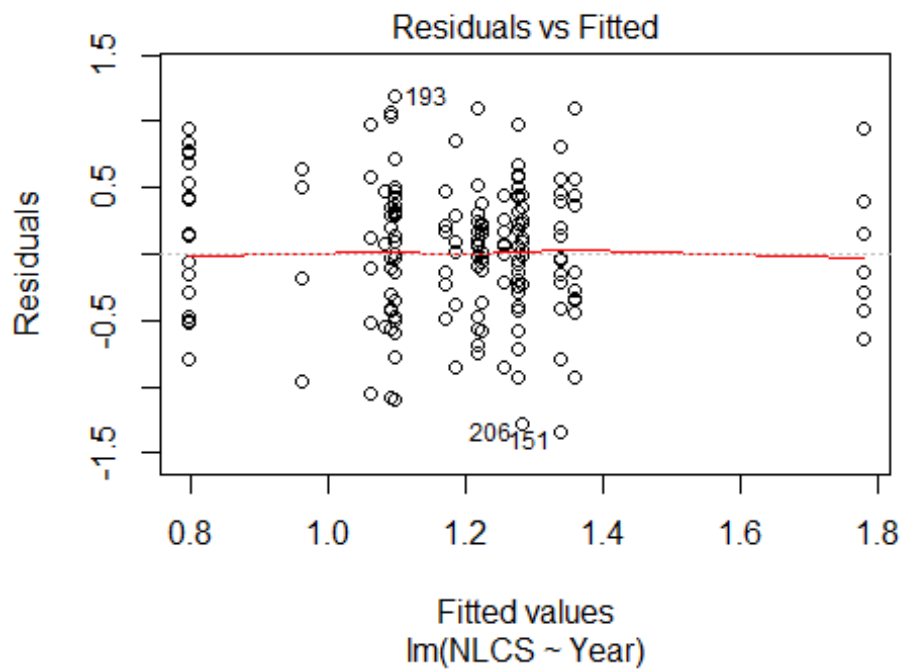
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.513 -0.291 0.037 0.346 0.984
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9870 0.6207 1.59 0.118
## LastAuthorFemale1 0.3660 0.1591 2.30 0.025 *
## Year1997 -0.9870 0.6207 -1.59 0.118
## Year1998 0.0910 0.8136 0.11 0.911
## Year1999 0.3470 0.6207 0.56 0.579
## Year2000 -0.0499 0.7468 -0.07 0.947
## Year2001 -0.3195 0.6335 -0.50 0.616
## Year2002 0.4365 0.6168 0.71 0.482
## Year2003 -0.3955 0.6782 -0.58 0.562
## Year2004 -0.0678 0.6655 -0.10 0.919
## Year2005 0.0535 0.6244 0.09 0.932
## Year2006 -0.2949 0.6663 -0.44 0.660
```

```

## Year2007      -0.6150      0.6307     -0.98      0.334
## Year2008      0.2271      0.6419      0.35      0.725
## Year2009     -0.4741      0.6556     -0.72      0.473
## Year2010      0.1604      1.0151      0.16      0.875
## Year2011      0.0320      0.6519      0.05      0.961
## Year2012     -0.3300      0.6391     -0.52      0.608
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.655
## Multiple R-squared:  0.273, Adjusted R-squared:  0.0356
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 59 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.573  0.910  0.955  0.935  0.986  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      1.43e-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: 70"
## [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
##    6    6    3    8   13   27   13   12    6    8    9   14   10   17   27
## 2011 2012
##   13   18
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    6    6    3    6   10   22   11   12    4    7    9   14   10   14   24
## 2011 2012

```

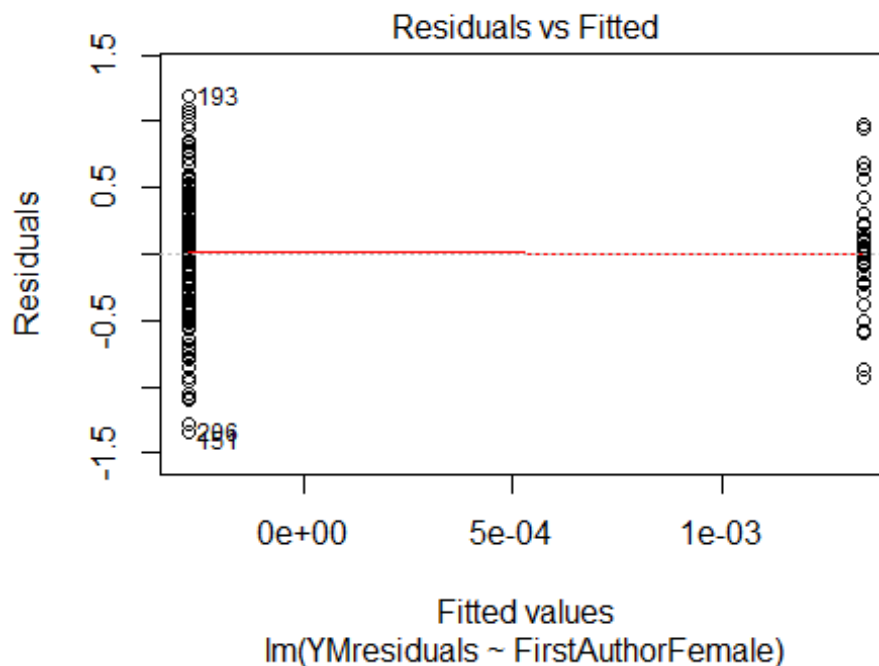
```
## 10 17
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 6 6 3 6 10 20 11 10 4 7 9 14 10 13 20
## 2011 2012
## 9 16
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 13, df = 16, p-value = 0.6
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.76, df = 1, p-value = 0.4
## [1] "Female first author team size 2018 geometric mean: 1.36242259325121"
## [1] "Male first author team size 2018 geometric mean: 2.82842712474619"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

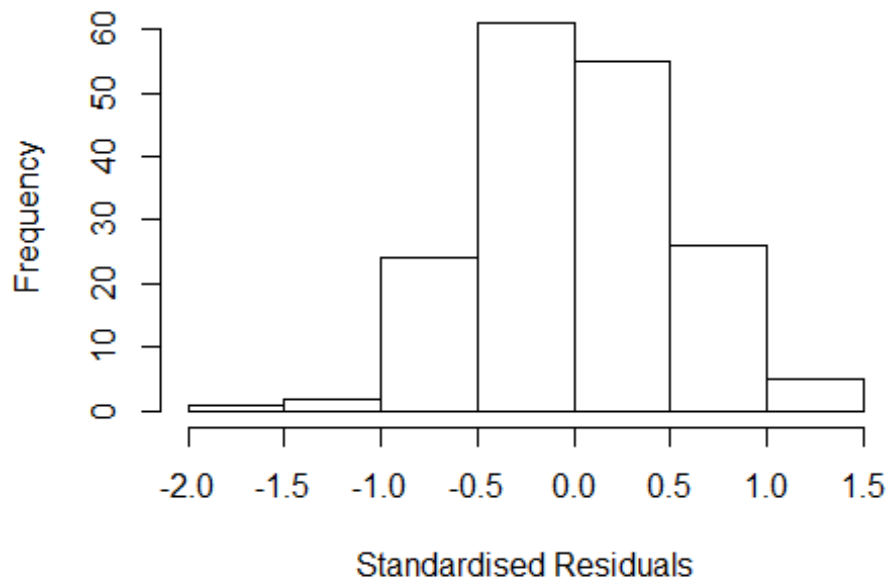
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 4.5, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male last author team size 2018 geometric mean: 3"

## 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 = 2, 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  2.47  1      1.572
## LastAuthorFemale  2.56  1      1.600
## UniqueAuthors    25.54  4      1.499
## Year              47.22 16      1.128
```


Residuals from first and last author and team size



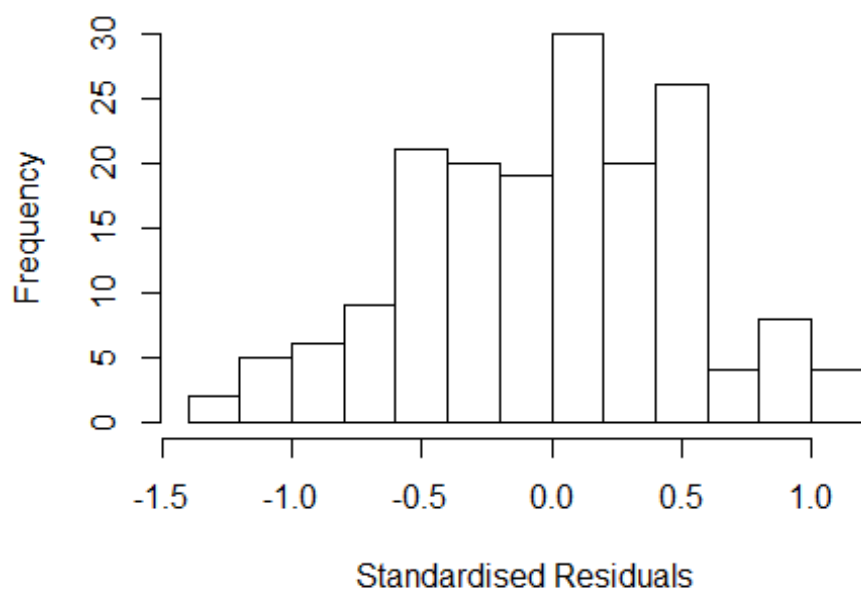
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.51006 -0.30705 -0.00587 0.30964 1.20465
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.00e+00 2.49e-01 4.03 8.9e-05 ***
## FirstAuthorFemale1 -2.87e-02 1.18e-01 -0.24 0.8084
## LastAuthorFemale1 1.27e-01 1.17e-01 1.09 0.2778
## UniqueAuthors2 3.05e-01 1.07e-01 2.85 0.0049 **
## UniqueAuthors3 3.04e-01 1.13e-01 2.68 0.0081 **
## UniqueAuthors4 3.83e-01 1.92e-01 1.99 0.0480 *
## UniqueAuthors5 1.35e-01 1.92e-01 0.70 0.4829
## Year1997 -1.60e-01 4.53e-01 -0.35 0.7250
## Year1998 -1.92e-01 4.02e-01 -0.48 0.6336
## Year1999 -8.73e-05 2.62e-01 0.00 0.9997
```

```

## Year2000          1.98e-01   3.10e-01    0.64   0.5231
## Year2001          -3.54e-01   2.98e-01   -1.19   0.2366
## Year2002           1.03e-01   2.45e-01    0.42   0.6752
## Year2003          -4.12e-02   3.56e-01   -0.12   0.9080
## Year2004          -1.39e-01   4.26e-01   -0.33   0.7450
## Year2005           6.44e-01   3.30e-01    1.96   0.0524 .
## Year2006           8.42e-02   2.62e-01    0.32   0.7487
## Year2007           1.46e-02   2.73e-01    0.05   0.9573
## Year2008           2.18e-01   3.20e-01    0.68   0.4976
## Year2009           1.35e-01   3.00e-01    0.45   0.6536
## Year2010          -2.20e-01   2.54e-01   -0.87   0.3884
## Year2011           1.03e-01   2.54e-01    0.41   0.6856
## Year2012           6.36e-02   2.52e-01    0.25   0.8012
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.19,   Adjusted R-squared:  0.0714
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 163 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.358  0.878  0.957   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          5.75e-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 2.023  1          1.422
## LastAuthorFemale  2.027  1          1.424
## Year              2.336 16          1.027

```

Residuals from first and last author



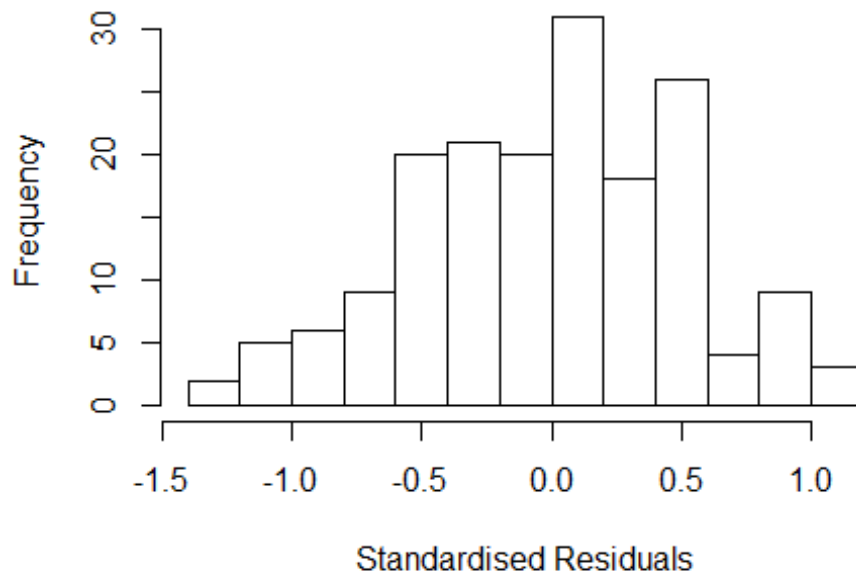
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.3601 -0.3958 0.0336 0.3848 1.1264
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.18891 0.24462 4.86 2.8e-06 ***
## FirstAuthorFemale1 -0.04180 0.12306 -0.34 0.735
## LastAuthorFemale1 0.05209 0.12277 0.42 0.672
## Year1997 -0.12614 0.41247 -0.31 0.760
## Year1998 -0.09531 0.34831 -0.27 0.785
## Year1999 -0.01707 0.27777 -0.06 0.951
## Year2000 0.13337 0.32099 0.42 0.678
## Year2001 -0.42828 0.31031 -1.38 0.170
## Year2002 0.03061 0.25702 0.12 0.905
## Year2003 -0.06602 0.34581 -0.19 0.849
## Year2004 -0.17705 0.46723 -0.38 0.705
## Year2005 0.54762 0.32981 1.66 0.099 .
```

```

## Year2006          0.07995      0.26614      0.30      0.764
## Year2007          0.00269      0.28047      0.01      0.992
## Year2008          0.11551      0.32547      0.35      0.723
## Year2009          0.16088      0.28664      0.56      0.575
## Year2010         -0.18942      0.27437     -0.69      0.491
## Year2011          0.15130      0.27016      0.56      0.576
## Year2012          0.08873      0.26866      0.33      0.742
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.519
## Multiple R-squared:  0.139, Adjusted R-squared:  0.0394
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 159 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.472  0.887   0.943   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      5.75e-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 1.607 1          1.268
## Year              1.607 16          1.015

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.3562 -0.3872 0.0446 0.3838 1.1277
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.19089 0.24096 4.94 2e-06 ***
## FirstAuthorFemale1 -0.00613 0.11302 -0.05 0.957
## Year1997 -0.11284 0.41153 -0.27 0.784
## Year1998 -0.09589 0.35739 -0.27 0.789
## Year1999 -0.01485 0.27119 -0.05 0.956
## Year2000 0.14653 0.30987 0.47 0.637
## Year2001 -0.41477 0.29499 -1.41 0.162
## Year2002 0.04007 0.24725 0.16 0.871
## Year2003 -0.05147 0.33588 -0.15 0.878
## Year2004 -0.16770 0.45883 -0.37 0.715
## Year2005 0.56186 0.31618 1.78 0.078 .
## Year2006 0.08959 0.25800 0.35 0.729
```

```

## Year2007          0.01765    0.26899    0.07    0.948
## Year2008          0.12914    0.31576    0.41    0.683
## Year2009          0.17144    0.27988    0.61    0.541
## Year2010         -0.18218    0.26836   -0.68    0.498
## Year2011          0.15841    0.26608    0.60    0.552
## Year2012          0.09778    0.26300    0.37    0.711
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.521
## Multiple R-squared:  0.138, Adjusted R-squared:  0.0443
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 164 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.478  0.891  0.945  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      5.75e-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 1.65  1      1.285
## Year              1.65 16      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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.3645 -0.3962  0.0419  0.3830  1.1255

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1713    0.2352   4.98 1.7e-06 ***
## LastAuthorFemale1  0.0264    0.1115   0.24  0.813
## Year1997         -0.1182    0.4133  -0.29  0.775
## Year1998         -0.1011    0.3540  -0.29  0.776
## Year1999         -0.0173    0.2761  -0.06  0.950
## Year2000          0.1357    0.3215   0.42  0.673
## Year2001         -0.4270    0.3097  -1.38  0.170
## Year2002          0.0381    0.2539   0.15  0.881
## Year2003         -0.0648    0.3455  -0.19  0.852
## Year2004         -0.1697    0.4718  -0.36  0.719
## Year2005          0.5488    0.3310   1.66  0.099 .
## Year2006          0.0864    0.2645   0.33  0.744
## Year2007          0.0105    0.2772   0.04  0.970
## Year2008          0.1183    0.3245   0.36  0.716
## Year2009          0.1667    0.2841   0.59  0.558
## Year2010         -0.1855    0.2727  -0.68  0.497
## Year2011          0.1606    0.2665   0.60  0.548
## Year2012          0.0908    0.2683   0.34  0.735
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.514
## Multiple R-squared:  0.139, Adjusted R-squared:  0.0452
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 161 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.462  0.891  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      5.75e-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: 174"
## [1] ""

```

```

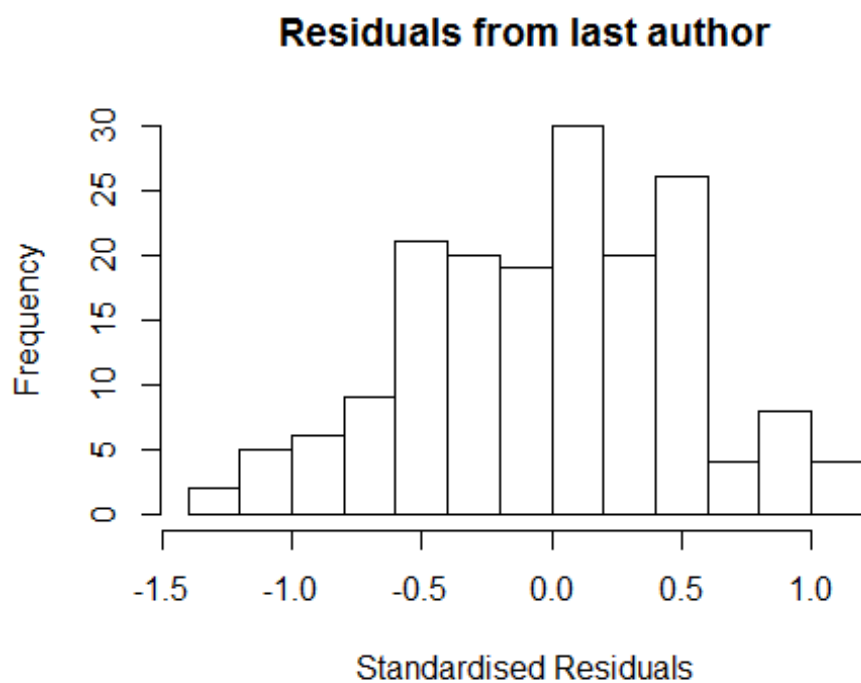
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3319"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    9    3    3    3    5    5    2    6    6    7   10   15    3    8
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    3    1    3    3    5    5    2    3    6    6    8   13    3    6
##
## 1997 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    3    1    3    2    5    5    2    3    5    6    8   10    3    6
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.59455793396046"
## [1] "Male first author team size 2018 geometric mean: 2.64575131106459"

## 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 = 9.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.93432994795422"
## [1] "Male last author team size 2018 geometric mean: 1.81712059283214"

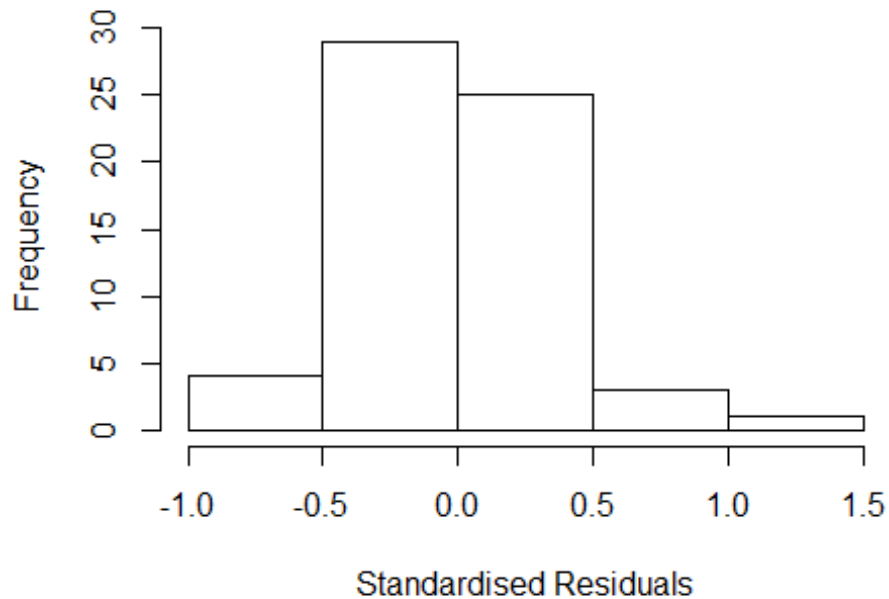
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```

```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 20, 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  9.36  1      3.059
## LastAuthorFemale   5.54  1      2.354
## UniqueAuthors    121.46  4      1.822
## Year              535.74 13      1.273
```

Residuals from first and last author and team size



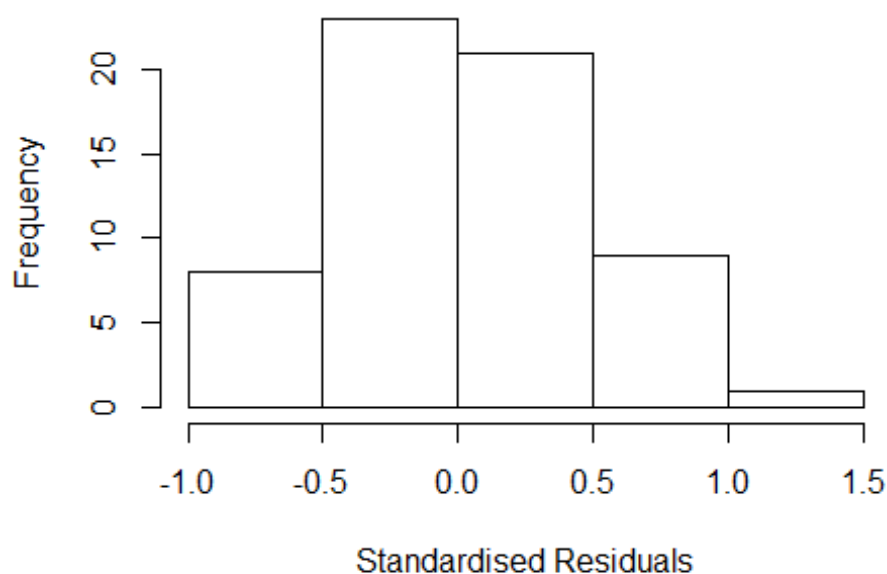
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9174 -0.3155 -0.0149 0.3044 1.0979
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8368 0.2795 2.99 0.0046 **
## FirstAuthorFemale1 -0.0658 0.1337 -0.49 0.6251
## LastAuthorFemale1 -0.0863 0.1269 -0.68 0.5001
## UniqueAuthors2 0.3404 0.1584 2.15 0.0375 *
## UniqueAuthors3 0.2571 0.2290 1.12 0.2679
## UniqueAuthors4 0.9219 0.1654 5.57 1.6e-06 ***
## UniqueAuthors5 -0.2282 0.2964 -0.77 0.4457
## Year2000 0.0632 0.2795 0.23 0.8223
## Year2001 0.0701 0.3491 0.20 0.8419
## Year2002 -0.2471 0.3477 -0.71 0.4812
```

```

## Year2003          0.0672      0.3419      0.20      0.8451
## Year2004          0.0303      0.3194      0.09      0.9248
## Year2005         -0.2253      0.4625     -0.49      0.6287
## Year2006         -0.0395      0.4568     -0.09      0.9315
## Year2007         -0.4648      0.2700     -1.72      0.0925 .
## Year2008          0.2704      0.3635      0.74      0.4610
## Year2009          0.0828      0.2652      0.31      0.7566
## Year2010          0.2313      0.3621      0.64      0.5264
## Year2011          0.2847      0.3026      0.94      0.3522
## Year2012         -0.1262      0.2688     -0.47      0.6411
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.509
## Multiple R-squared:  0.412, Adjusted R-squared:  0.146
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 56 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.621  0.928  0.962  0.938  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      1.61e-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 4.893 1      2.212
## LastAuthorFemale 4.805 1      2.192
## Year      13.150 13      1.104

```

Residuals from first and last author



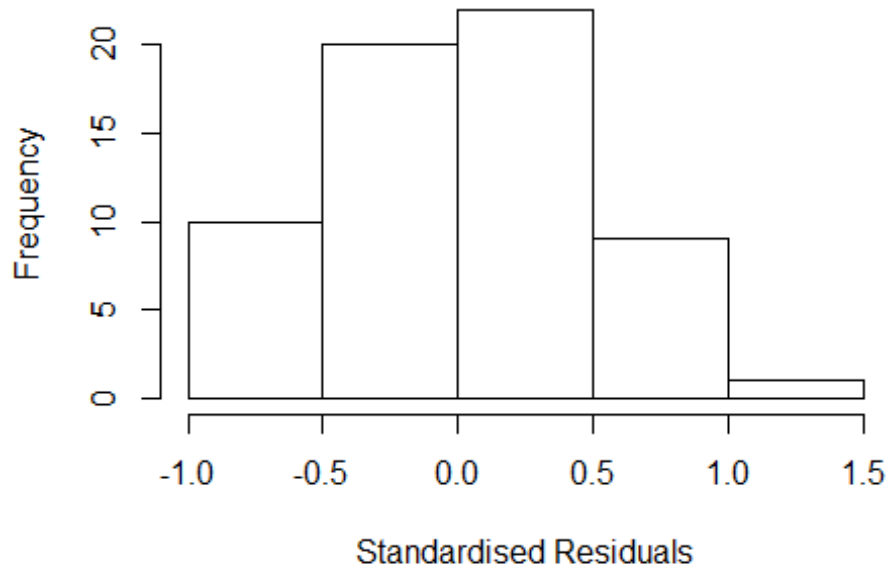
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.94095 -0.32268 0.00629 0.39979 1.28520
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.95426 0.24082 3.96 0.00026 ***
## FirstAuthorFemale1 -0.07728 0.13533 -0.57 0.57075
## LastAuthorFemale1 -0.11200 0.15106 -0.74 0.46219
## Year2000 -0.05426 0.24082 -0.23 0.82274
## Year2001 0.36565 0.37155 0.98 0.33021
## Year2002 -0.18862 0.26643 -0.71 0.48254
## Year2003 0.17597 0.35628 0.49 0.62372
## Year2004 0.00467 0.31274 0.01 0.98814
## Year2005 -0.14112 0.63152 -0.22 0.82417
## Year2006 0.19677 0.41806 0.47 0.64010
## Year2007 -0.44230 0.30984 -1.43 0.16018
## Year2008 0.47015 0.29428 1.60 0.11697
```

```

## Year2009          0.25091    0.26731    0.94  0.35282
## Year2010          0.22082    0.31495    0.70  0.48676
## Year2011          0.59668    0.36000    1.66  0.10424
## Year2012         -0.14008    0.27943   -0.50  0.61856
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.546
## Multiple R-squared:  0.261, Adjusted R-squared:  0.0199
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 55 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.559  0.899  0.941  0.931  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      1.61e-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 5.526 1          2.351
## Year              5.526 13          1.068

```

Residuals from first author



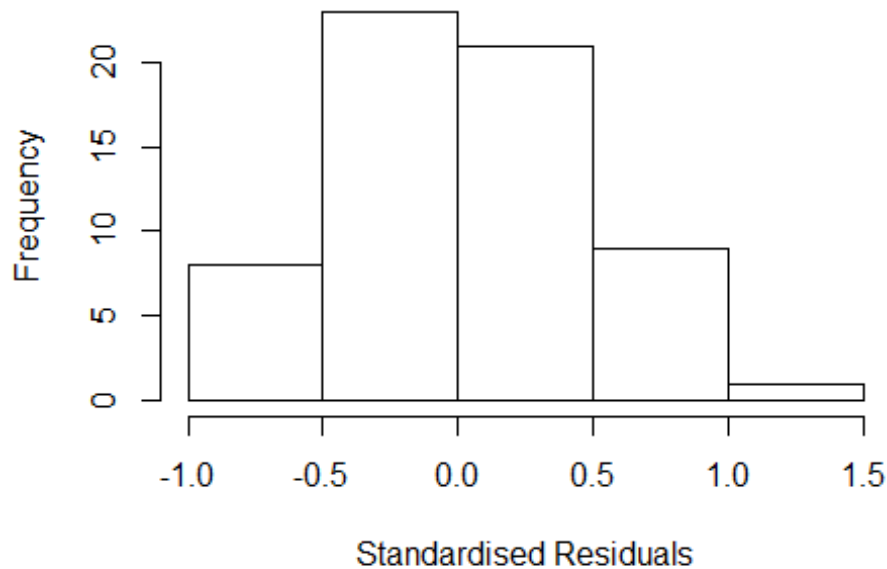
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9547 -0.3219 0.0259 0.4029 1.2819
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.89518 0.22668 3.95 0.00026 ***
## FirstAuthorFemale1 -0.13185 0.12719 -1.04 0.30518
## Year2000 0.00482 0.22668 0.02 0.98312
## Year2001 0.40649 0.35359 1.15 0.25613
## Year2002 -0.10225 0.23673 -0.43 0.66776
## Year2003 0.19138 0.36166 0.53 0.59918
## Year2004 0.04005 0.32358 0.12 0.90203
## Year2005 -0.16675 0.58943 -0.28 0.77849
## Year2006 0.24281 0.38823 0.63 0.53471
## Year2007 -0.44147 0.31901 -1.38 0.17293
## Year2008 0.50864 0.28638 1.78 0.08219 .
## Year2009 0.29416 0.27040 1.09 0.28219
```

```

## Year2010          0.22574    0.32623    0.69  0.49235
## Year2011          0.62804    0.40400    1.55  0.12676
## Year2012         -0.14862    0.28652   -0.52  0.60640
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.557
## Multiple R-squared:  0.255, Adjusted R-squared:  0.0332
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 5 weights are ~= 1. The remaining 57 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.575  0.905  0.953  0.935  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      1.61e-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 4.083 1          2.021
## Year            4.083 13          1.056

```

Residuals from last author



```
## [1] "List of 0 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
## -0.94798 -0.34325 0.00867 0.39707 1.26636
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9611 0.2490 3.86 0.00034 ***
## LastAuthorFemale1 -0.1570 0.1412 -1.11 0.27200
## Year2000 -0.0611 0.2490 -0.25 0.80720
## Year2001 0.3485 0.3668 0.95 0.34693
## Year2002 -0.2341 0.2652 -0.88 0.38188
## Year2003 0.1438 0.3614 0.40 0.69242
## Year2004 -0.0356 0.3083 -0.12 0.90862
## Year2005 -0.1416 0.6900 -0.21 0.83825
## Year2006 0.1833 0.4137 0.44 0.65978
## Year2007 -0.4609 0.3255 -1.42 0.16338
## Year2008 0.4327 0.2843 1.52 0.13473
## Year2009 0.2199 0.2584 0.85 0.39918
```

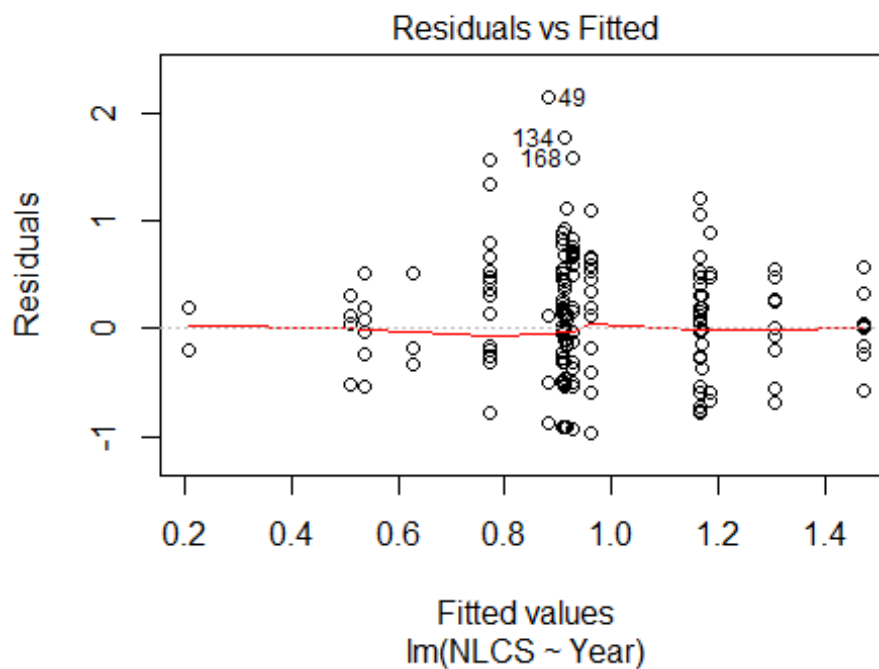


```

## Year2010          0.2005      0.3330      0.60  0.55004
## Year2011          0.5460      0.3498      1.56  0.12521
## Year2012         -0.1651      0.2816     -0.59  0.56038
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.557
## Multiple R-squared:  0.257, Adjusted R-squared:  0.0355
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 56 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.585  0.905  0.950  0.935  0.991  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      1.61e-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: 62"
## [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
##    2    6    6    2    7    8   11    9    7   10   12   20   27   29   31
## 2011 2012
##   35   27
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    0    5    3    2    6    6    8    5    5    9   10   18   20   25   29
## 2011 2012
##   26   23
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010

```

```
##      0      5      3      2      5      6      8      5      5      9     10     18     19     23     29
## 2011 2012
##    25    22
## [1] "Heteroscedasticity checks, confirming that there are problems with
##      these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 25, df = 15, p-value = 0.05
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.55, df = 1, p-value = 0.5

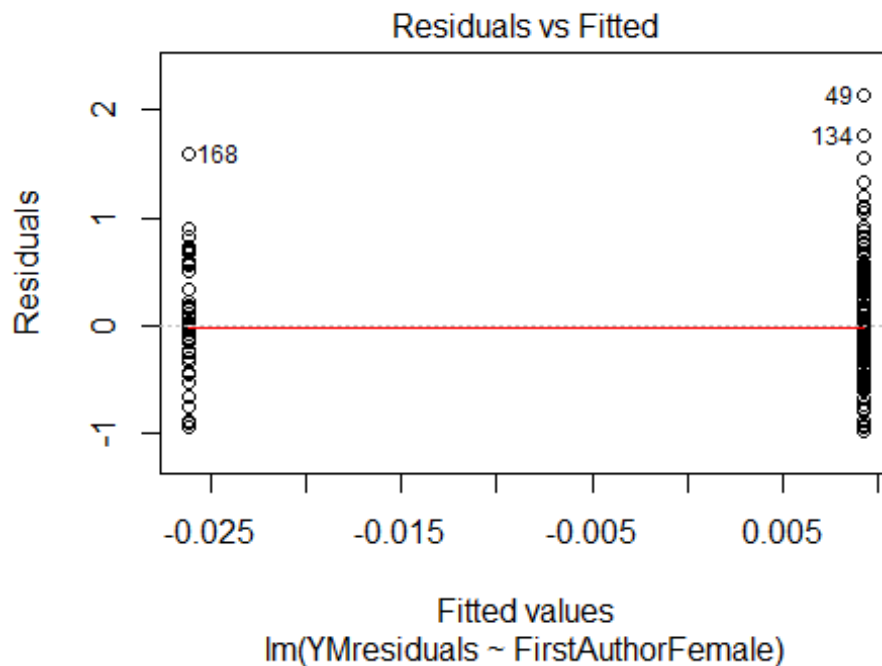
## [1] "Female first author team size 2018 geometric mean: 1.14869835499704"
## [1] "Male first author team size 2018 geometric mean: 1.08005973889231"

## 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 = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 1.25992104989487"
## [1] "Male last author team size 2018 geometric mean: 1.02811382665607"

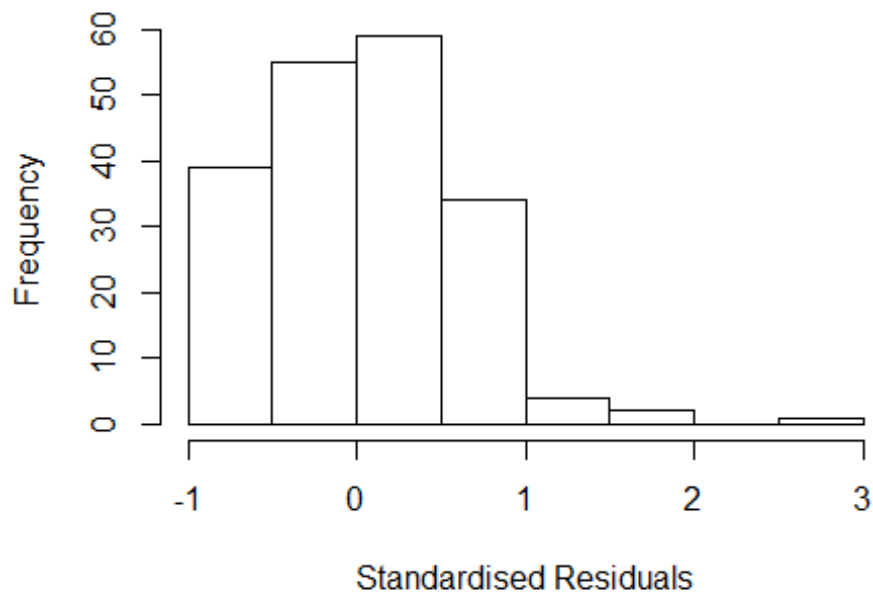
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 180, 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	5.460	1	2.337
LastAuthorFemale	5.013	1	2.239
UniqueAuthors	2.952	3	1.198
Year	4.682	15	1.053

Residuals from first and last author and team size



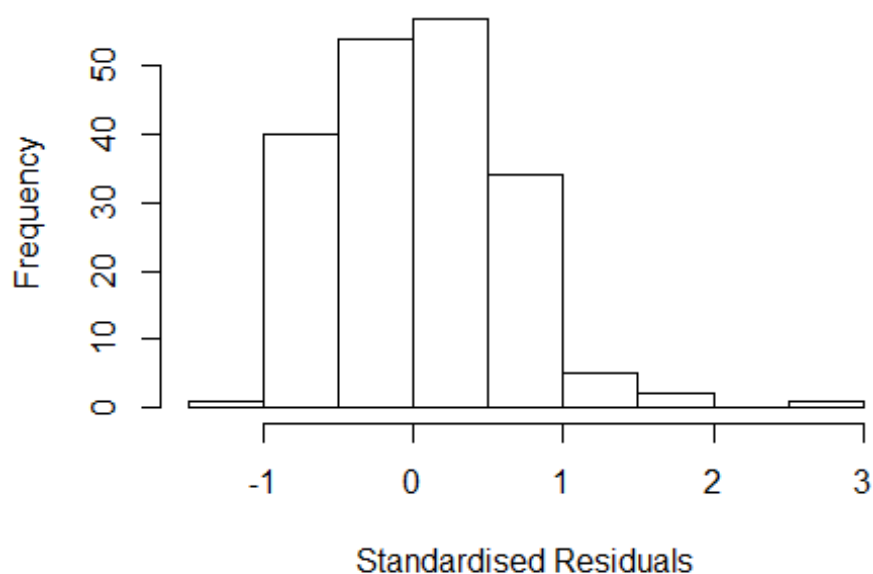
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 49 0042745639 3.028 2003    3320      1    2.698
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.94296 -0.38812  0.00762  0.40056  2.69763
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.15131    0.12272   9.38 < 2e-16 ***
## FirstAuthorFemale1 -0.16878    0.19636  -0.86  0.39122
## LastAuthorFemale1  0.17910    0.18416   0.97  0.33215
## UniqueAuthors2     0.04731    0.11742   0.40  0.68751
## UniqueAuthors3     0.70643    0.17231   4.10 6.4e-05 ***
## UniqueAuthors4     0.00796    0.13051   0.06  0.95141
## Year1998          -0.53291    0.25774  -2.07  0.04016 *
## Year1999          -0.95096    0.19856  -4.79 3.6e-06 ***
## Year2000           0.15074    0.34527   0.44  0.66296
## Year2001          -0.61701    0.20106  -3.07  0.00250 **
```

```

## Year2002          0.30704      0.16677      1.84  0.06733 .
## Year2003         -0.82094      0.24778     -3.31  0.00112 **
## Year2004         -0.63629      0.17498     -3.64  0.00036 ***
## Year2005         -0.26899      0.19210     -1.40  0.16323
## Year2006          0.07582      0.16628      0.46  0.64896
## Year2007         -0.21867      0.19753     -1.11  0.26984
## Year2008         -0.35402      0.18610     -1.90  0.05879 .
## Year2009         -0.35567      0.22237     -1.60  0.11154
## Year2010         -0.00554      0.15225     -0.04  0.97102
## Year2011         -0.25366      0.15881     -1.60  0.11204
## Year2012         -0.49254      0.20341     -2.42  0.01649 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.556
## Multiple R-squared:  0.22,   Adjusted R-squared:  0.13
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## observation 30 is an outlier with |weight| = 0 ( < 0.00052);
## 23 weights are ~= 1. The remaining 170 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.330  0.871  0.940  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          5.15e-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 4.628 1          2.151
## LastAuthorFemale  4.275 1          2.068
## Year              1.704 15          1.018

```

Residuals from first and last author



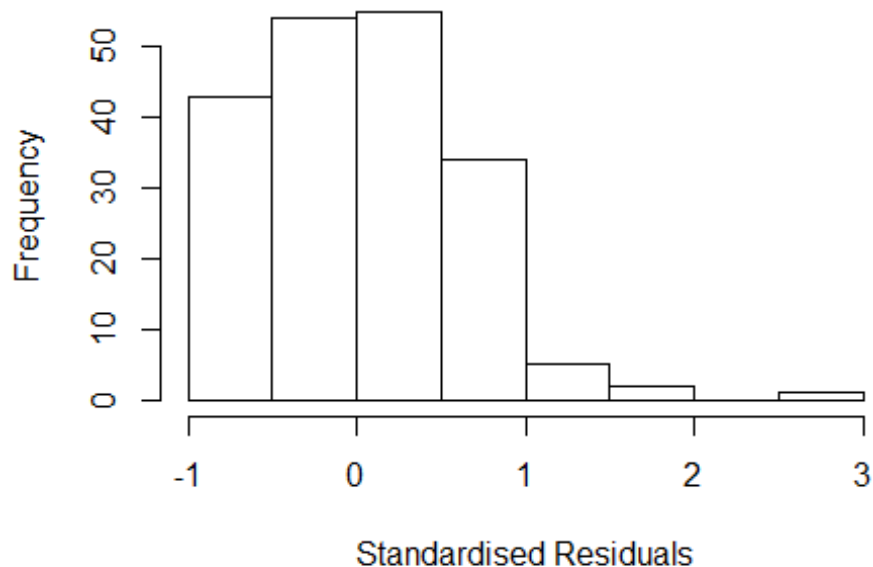
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 49 0042745639 3.028 2003    3320      1    2.696
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.00039 -0.38586  0.00929  0.44616  2.69607
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1717    0.1100   10.65 < 2e-16 ***
## FirstAuthorFemale1 -0.1357    0.1871   -0.72  0.46948
## LastAuthorFemale1  0.1745    0.1758    0.99  0.32226
## Year1998         -0.5524    0.2514   -2.20  0.02928 *
## Year1999         -0.9856    0.1811   -5.44  1.7e-07 ***
## Year2000          0.1378    0.3432    0.40  0.68847
## Year2001         -0.6529    0.1899   -3.44  0.00073 ***
## Year2002          0.2901    0.1641    1.77  0.07879 .
## Year2003         -0.8397    0.2451   -3.43  0.00076 ***
## Year2004         -0.6628    0.1631   -4.06  7.3e-05 ***
## Year2005         -0.2941    0.1940   -1.52  0.13137
## Year2006          0.1265    0.1724    0.73  0.46402
```

```

## Year2007          -0.2101      0.2034   -1.03   0.30288
## Year2008          -0.3795      0.1781   -2.13   0.03453 *
## Year2009          -0.2816      0.2173   -1.30   0.19661
## Year2010          -0.0239      0.1516   -0.16   0.87476
## Year2011          -0.2399      0.1564   -1.53   0.12689
## Year2012          -0.5080      0.2035   -2.50   0.01347 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.184, Adjusted R-squared:  0.105
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 30 is an outlier with |weight| <= 0.00019 ( < 0.00052);
## 22 weights are ~= 1. The remaining 171 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.381  0.871  0.935  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      5.15e-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 1.353 1          1.163
## Year              1.353 15          1.010

```

Residuals from first author



```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 49 0042745639 3.028 2003    3320      1    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
## -0.98923 -0.36191  0.00288  0.43520  2.69652
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1717    0.1101   10.64 < 2e-16 ***
## FirstAuthorFemale1  0.0164    0.1011    0.16  0.87154
## Year1998        -0.5526    0.2516   -2.20  0.02938 *
## Year1999        -0.9744    0.1871   -5.21  5.3e-07 ***
## Year2000         0.1705    0.3717    0.46  0.64703
## Year2001        -0.6407    0.1903   -3.37  0.00093 ***
## Year2002         0.2962    0.1640    1.81  0.07255 .
## Year2003        -0.8402    0.2417   -3.48  0.00064 ***
## Year2004        -0.6583    0.1647   -4.00  9.4e-05 ***
## Year2005        -0.2843    0.1918   -1.48  0.14002
## Year2006         0.1354    0.1719    0.79  0.43201
## Year2007        -0.1988    0.2002   -0.99  0.32195
```

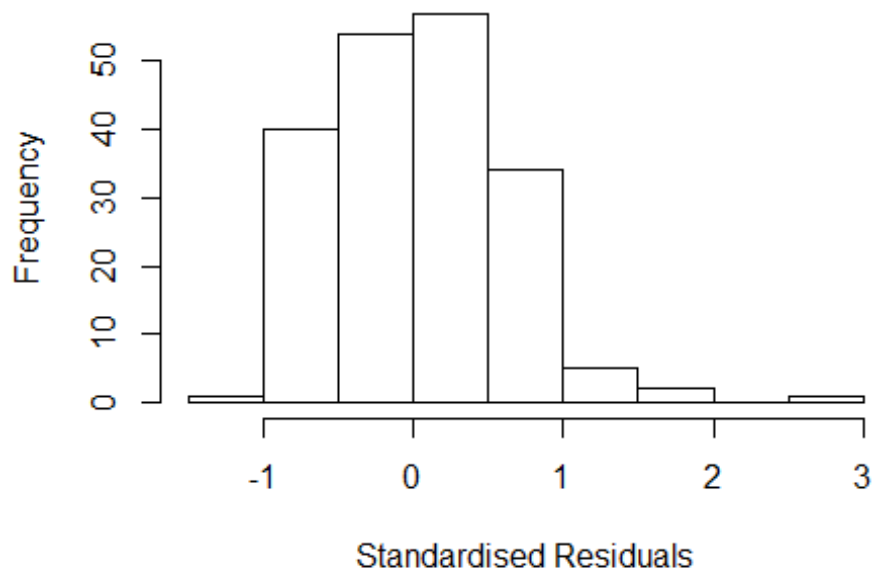


```

## Year2008          -0.3711      0.1783   -2.08  0.03885 *
## Year2009          -0.2497      0.2140   -1.17  0.24480
## Year2010          -0.0200      0.1516   -0.13  0.89495
## Year2011          -0.2282      0.1578   -1.45  0.14985
## Year2012          -0.5134      0.2011   -2.55  0.01152 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.574
## Multiple R-squared:  0.18,   Adjusted R-squared:  0.106
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## observation 30 is an outlier with |weight| = 0 ( < 0.00052);
## 18 weights are ~= 1. The remaining 175 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.369  0.878  0.936  0.914  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      5.15e-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 1.281 1          1.132
## Year            1.281 15          1.008

```

Residuals from last author



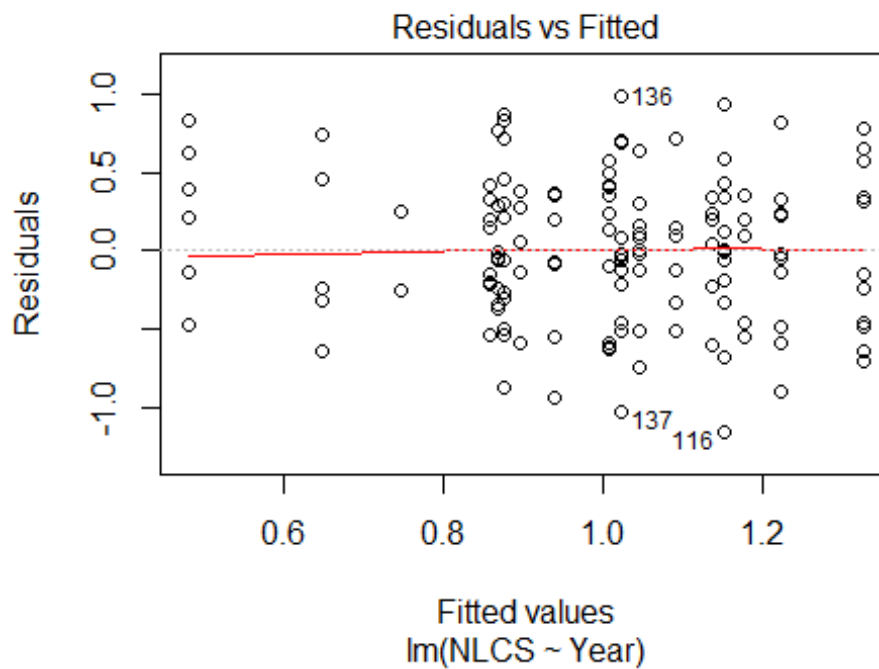
```
## [1] "List of 1 outliers with residuals above 2.5"
##      ScopusId  NLCS Year OneField Fields residuals
## 49 0042745639 3.028 2003    3320      1    2.696
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
## AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.02408 -0.37418  0.00431  0.43903  2.69639
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1717    0.1101   10.64 < 2e-16 ***
## LastAuthorFemale1  0.0661    0.0966    0.68  0.49486
## Year1998        -0.5525    0.2515   -2.20  0.02933 *
## Year1999        -0.9992    0.1727   -5.79  3.2e-08 ***
## Year2000         0.1587    0.3596    0.44  0.65941
## Year2001        -0.6676    0.1871   -3.57  0.00046 ***
## Year2002         0.2826    0.1636    1.73  0.08586 .
## Year2003        -0.8401    0.2421   -3.47  0.00066 ***
## Year2004        -0.6682    0.1612   -4.15  5.2e-05 ***
## Year2005        -0.3067    0.1952   -1.57  0.11798
## Year2006         0.1159    0.1725    0.67  0.50230
## Year2007        -0.2137    0.2013   -1.06  0.28982
```

```

## Year2008          -0.3903      0.1772    -2.20  0.02889 *
## Year2009          -0.2660      0.2144    -1.24  0.21632
## Year2010          -0.0291      0.1515    -0.19  0.84788
## Year2011          -0.2419      0.1565    -1.55  0.12395
## Year2012          -0.5278      0.1994    -2.65  0.00885 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.576
## Multiple R-squared:  0.182, Adjusted R-squared:  0.108
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## observation 30 is an outlier with |weight| <= 4.6e-06 ( < 0.00052);
## 21 weights are ~= 1. The remaining 172 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.365  0.871  0.935   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      5.15e-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: 194"
## [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
##    7    6    6   10   11   10   13    7    3   13    7   12   12   13    9
## 2011 2012
##   18   14
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    5    6    5    8    8    9   12    6    2   11    6   12   12   13    9
## 2011 2012

```

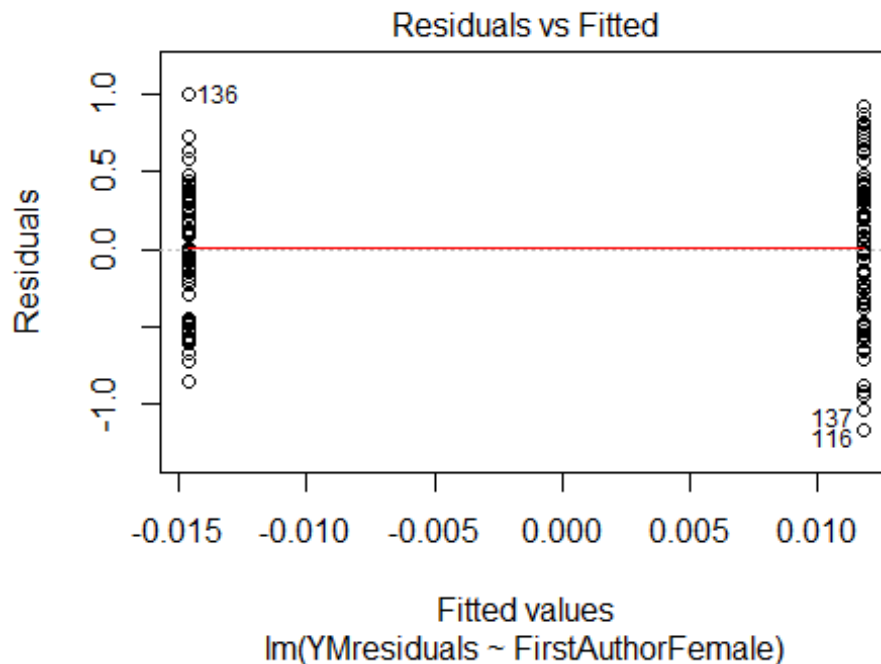
```
## 13 11
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
## 5 6 5 8 7 9 12 6 2 11 6 11 12 12 9
## 2011 2012
## 12 10
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data: NLCS by Year
## Bartlett's K-squared = 8.5, df = 16, p-value = 0.9
```



```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 3.7, df = 1, p-value = 0.05
## [1] "Female first author team size 2018 geometric mean: 2.297396709999407"
## [1] "Male first author team size 2018 geometric mean: 1.91947121957741"
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

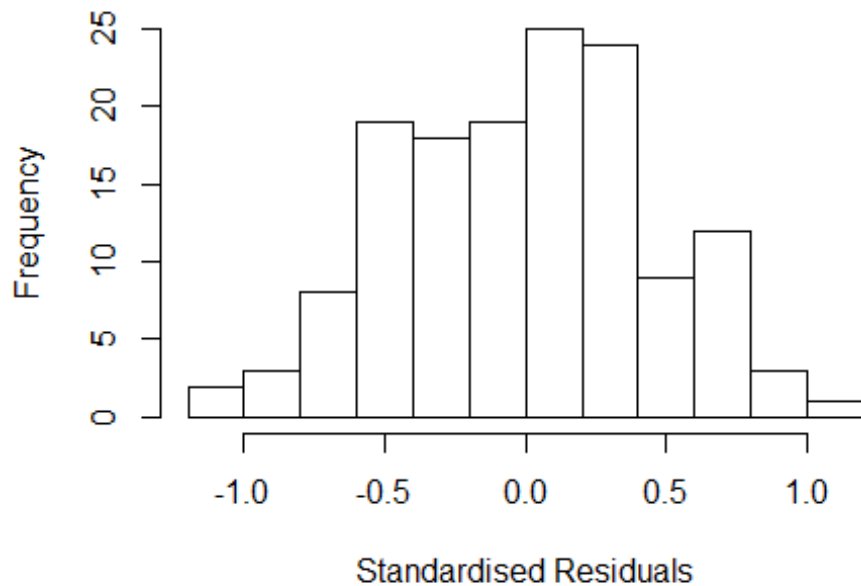
```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 21, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.29739670999407"
## [1] "Male last author team size 2018 geometric mean: 1.91947121957741"

## 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 = 21, 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  7.416  1          2.723
## LastAuthorFemale  4.814  1          2.194
## UniqueAuthors    164.139  4          1.892
## Year              173.290 16          1.175
```

Residuals from first and last author and team size



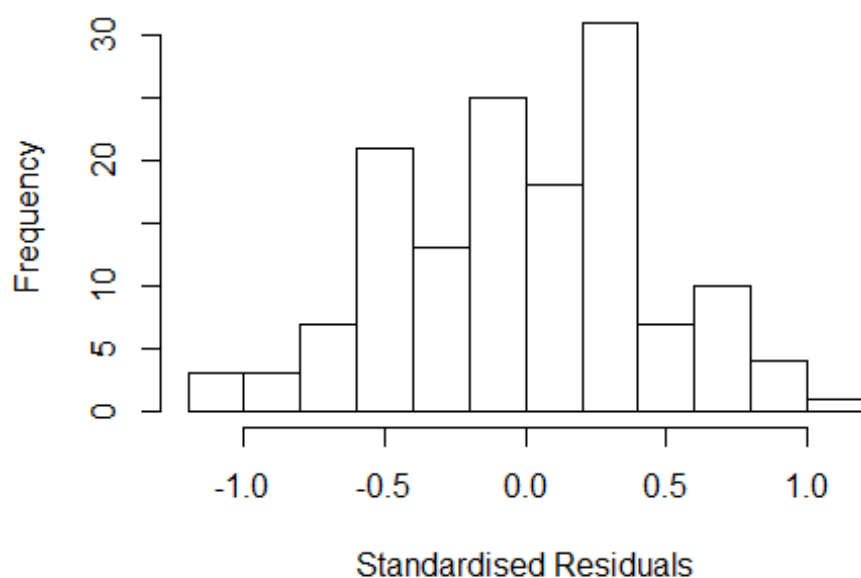
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
## lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.1502 -0.3459 0.0295 0.3142 1.0832
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.5893 0.2877 2.05 0.043 *
## FirstAuthorFemale1 -0.0180 0.1404 -0.13 0.898
## LastAuthorFemale1 -0.0555 0.1293 -0.43 0.669
## UniqueAuthors2 0.1000 0.1074 0.93 0.354
## UniqueAuthors3 0.1718 0.2090 0.82 0.413
## UniqueAuthors4 0.0967 0.2114 0.46 0.648
## UniqueAuthors5 0.4330 0.2342 1.85 0.067 .
## Year1997 0.5205 0.3013 1.73 0.087 .
## Year1998 0.3504 0.3310 1.06 0.292
## Year1999 0.2904 0.3082 0.94 0.348
```

```

## Year2000          0.3100      0.3042      1.02      0.310
## Year2001         -0.1389      0.3325     -0.42      0.677
## Year2002          0.6287      0.3136      2.00      0.047 *
## Year2003          0.5308      0.3388      1.57      0.120
## Year2004          0.1440      0.3821      0.38      0.707
## Year2005          0.4502      0.3078      1.46      0.146
## Year2006          0.4600      0.3269      1.41      0.162
## Year2007          0.5609      0.3342      1.68      0.096 .
## Year2008          0.4100      0.3223      1.27      0.206
## Year2009          0.2684      0.3617      0.74      0.460
## Year2010          0.3955      0.3176      1.25      0.215
## Year2011          0.4538      0.3134      1.45      0.150
## Year2012          0.6333      0.3406      1.86      0.065 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.494
## Multiple R-squared:  0.174, Adjusted R-squared:  0.0231
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 132 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.568  0.899  0.958  0.921  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      6.99e-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 2.866 1      1.693
## LastAuthorFemale 2.326 1      1.525
## Year              2.130 16      1.024

```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.19608 -0.36312 -0.00494 0.29789 1.04010
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.608 0.291 2.09 0.039 *
## FirstAuthorFemale1 0.039 0.124 0.31 0.754
## LastAuthorFemale1 -0.102 0.117 -0.87 0.385
## Year1997 0.543 0.314 1.73 0.086 .
## Year1998 0.329 0.333 0.99 0.326
## Year1999 0.270 0.311 0.87 0.387
## Year2000 0.295 0.310 0.95 0.342
## Year2001 -0.141 0.347 -0.41 0.685
## Year2002 0.658 0.316 2.09 0.039 *
## Year2003 0.599 0.332 1.80 0.074 .
## Year2004 0.171 0.358 0.48 0.635
## Year2005 0.491 0.310 1.58 0.116
```

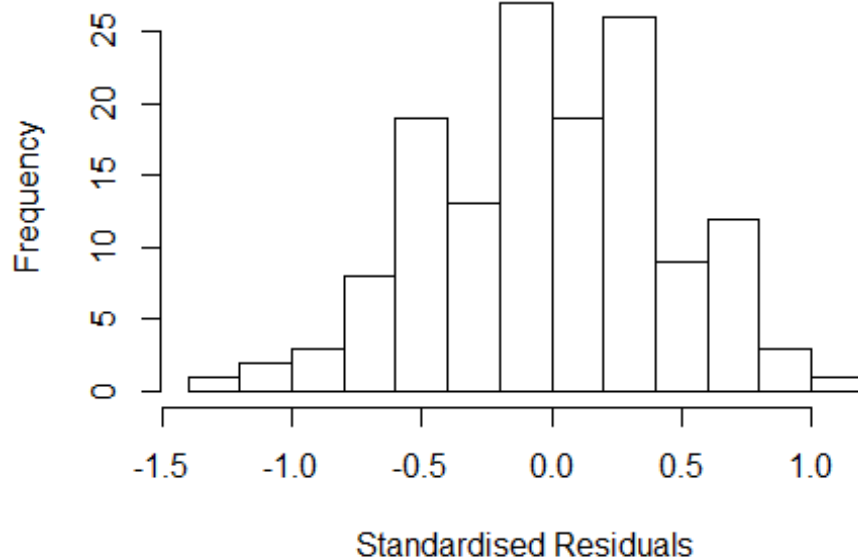


```

## Year2006          0.501      0.339      1.48      0.142
## Year2007          0.589      0.340      1.73      0.086 .
## Year2008          0.424      0.327      1.30      0.197
## Year2009          0.336      0.358      0.94      0.349
## Year2010          0.405      0.325      1.25      0.215
## Year2011          0.496      0.314      1.58      0.116
## Year2012          0.658      0.340      1.93      0.055 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.477
## Multiple R-squared:  0.164, Adjusted R-squared:  0.0423
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 14 weights are ~= 1. The remaining 129 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.509  0.881  0.946  0.913  0.981  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      6.99e-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 1.619 1      1.272
## Year              1.619 16      1.015

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.2056 -0.3546 -0.0102 0.3095 1.0107
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.6401 0.2915 2.20 0.030 *
## FirstAuthorFemale1 -0.0299 0.0936 -0.32 0.750
## Year1997 0.5186 0.3159 1.64 0.103
## Year1998 0.2817 0.3276 0.86 0.391
## Year1999 0.2286 0.3082 0.74 0.460
## Year2000 0.2677 0.3141 0.85 0.396
## Year2001 -0.1806 0.3473 -0.52 0.604
## Year2002 0.6120 0.3171 1.93 0.056 .
## Year2003 0.5668 0.3331 1.70 0.091 .
## Year2004 0.1213 0.3462 0.35 0.727
## Year2005 0.4309 0.3023 1.43 0.157
## Year2006 0.4520 0.3319 1.36 0.176
```

```

## Year2007          0.5655      0.3380      1.67      0.097 .
## Year2008          0.3880      0.3296      1.18      0.241
## Year2009          0.2849      0.3633      0.78      0.434
## Year2010          0.3627      0.3242      1.12      0.265
## Year2011          0.4580      0.3151      1.45      0.149
## Year2012          0.6009      0.3417      1.76      0.081 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.474
## Multiple R-squared:  0.16,   Adjusted R-squared:  0.0455
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 12 weights are ~= 1. The remaining 131 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.      Max.
##  0.497  0.869  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      6.99e-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 1.339 1      1.157
## Year      1.339 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.21284 -0.37453  0.00376  0.31514  1.03853

```

```

##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.6264    0.2794    2.24  0.027 *
## LastAuthorFemale1 -0.0754    0.0885   -0.85  0.396
## Year1997        0.5323    0.3074    1.73  0.086 .
## Year1998        0.3151    0.3289    0.96  0.340
## Year1999        0.2541    0.3038    0.84  0.404
## Year2000        0.2849    0.3088    0.92  0.358
## Year2001       -0.1573    0.3394   -0.46  0.644
## Year2002        0.6435    0.3126    2.06  0.042 *
## Year2003        0.5913    0.3343    1.77  0.079 .
## Year2004        0.1579    0.3572    0.44  0.659
## Year2005        0.4763    0.3029    1.57  0.118
## Year2006        0.4879    0.3360    1.45  0.149
## Year2007        0.5865    0.3406    1.72  0.088 .
## Year2008        0.4196    0.3298    1.27  0.206
## Year2009        0.3224    0.3552    0.91  0.366
## Year2010        0.3991    0.3256    1.23  0.223
## Year2011        0.4905    0.3128    1.57  0.119
## Year2012        0.6360    0.3305    1.92  0.057 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared:  0.163, Adjusted R-squared:  0.0491
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 130 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.500  0.880  0.946  0.914  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      6.99e-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: 143"
## [1] ""

```

```

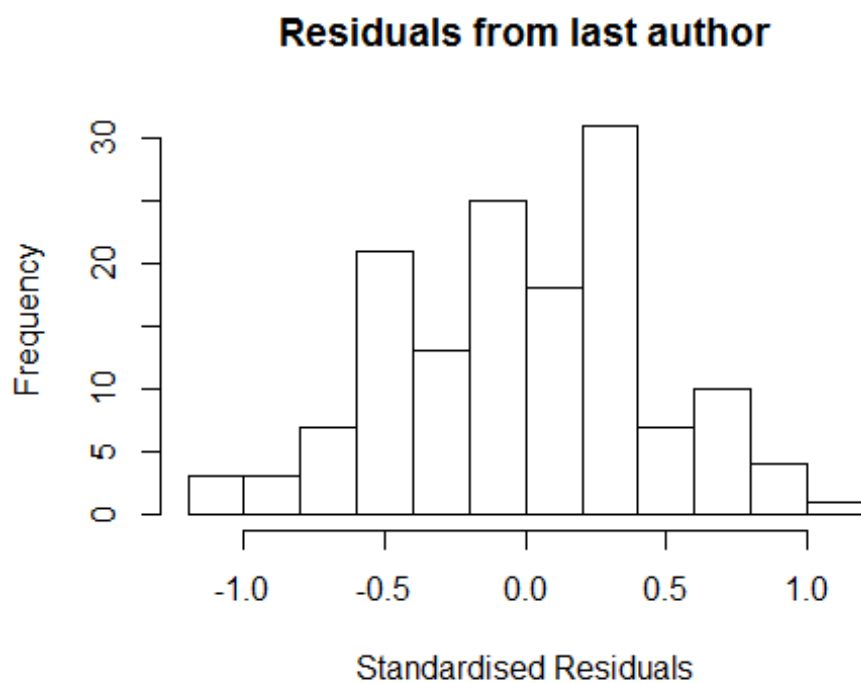
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3322"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    4    5    4    3    2    2    1    7    7    3    3    7    9    9    11
## 2011 2012
##   22    9
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    3    3    1    0    1    7    6    3    2    7    8    5    8
## 2011 2012
##   19    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    4    3    3    1    0    1    7    6    3    2    7    8    5    8
## 2011 2012
##   15    8
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 1.74110112659225"
## [1] "Male first author team size 2018 geometric mean: 1.57041780247502"

## 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] "Female last author team size 2018 geometric mean: 1.41421356237309"
## [1] "Male last author team size 2018 geometric mean: 1.70223744015616"

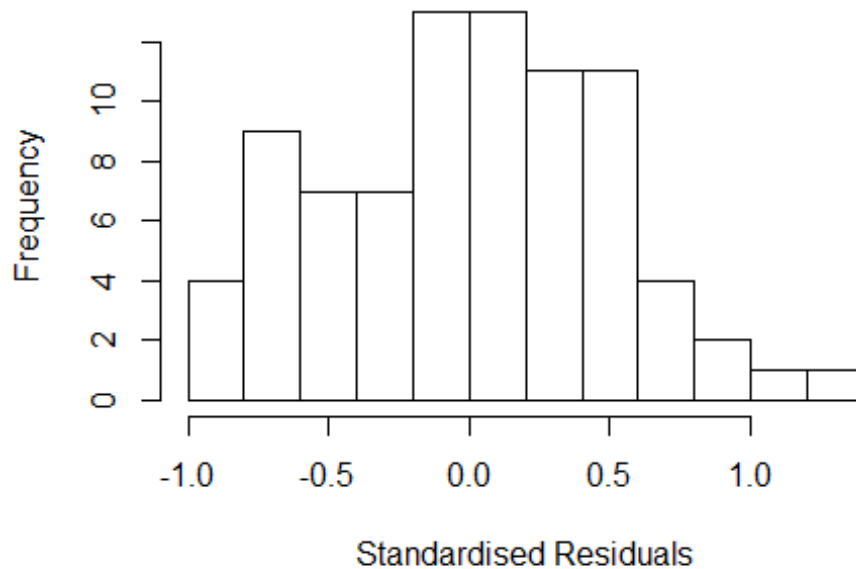
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 7.5, 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 1.228e+13  1      3.504e+06
## LastAuthorFemale  7.211e+02  1      2.685e+01
## UniqueAuthors    4.138e+16  4      1.194e+02
## Year              4.913e+17 15      3.888e+00
```

Residuals from first and last author and team size



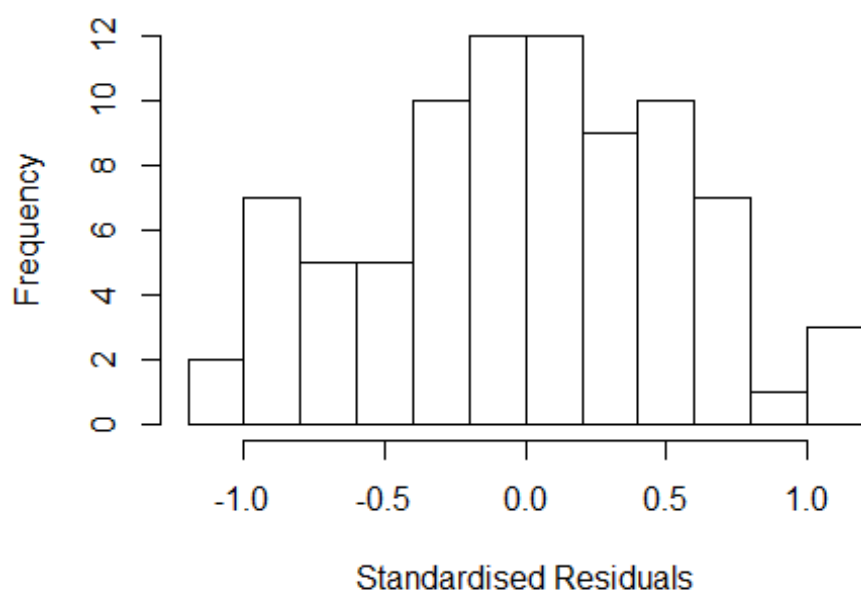
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.9654 -0.3309 0.0138 0.2880 1.2173
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1668 0.4359 2.68 0.0095 **
## FirstAuthorFemale1 0.0149 0.1986 0.07 0.9406
## LastAuthorFemale1 -0.1842 0.1765 -1.04 0.3009
## UniqueAuthors2 0.2370 0.1602 1.48 0.1443
## UniqueAuthors3 0.2895 0.1737 1.67 0.1006
## UniqueAuthors4 0.7819 0.2504 3.12 0.0027 **
## UniqueAuthors5 0.8103 0.1446 5.60 5.4e-07 ***
## Year1997 -0.1836 0.4883 -0.38 0.7082
## Year1998 -0.3986 0.5111 -0.78 0.4385
## Year1999 -0.1783 0.4361 -0.41 0.6842
```

```

## Year2000          -0.0358      0.4359   -0.08   0.9349
## Year2002           0.2309      0.4489    0.51   0.6089
## Year2003          -0.0424      0.4619   -0.09   0.9271
## Year2004          -0.2180      0.4982   -0.44   0.6632
## Year2005           0.0145      0.4291    0.03   0.9732
## Year2006           0.3914      0.8022    0.49   0.6273
## Year2007          -0.1407      0.4657   -0.30   0.7635
## Year2008           0.1029      0.4570    0.23   0.8226
## Year2009          -0.3766      0.4380   -0.86   0.3932
## Year2010          -0.2608      0.4346   -0.60   0.5507
## Year2011          -0.3711      0.4504   -0.82   0.4132
## Year2012          -0.5587      0.5111   -1.09   0.2786
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.254, Adjusted R-squared:  -0.00239
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 15 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.638  0.872  0.948  0.922  0.984  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      1.20e-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 2.220e+14 1      1.490e+07
## LastAuthorFemale  4.936e+00 1      2.222e+00
## Year              4.563e+14 15      3.081e+00

```


Residuals from first and last author



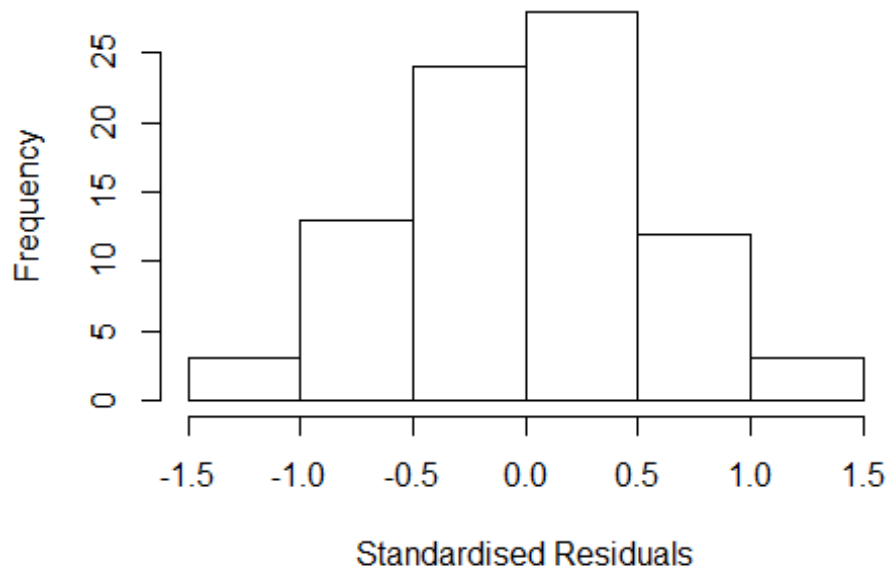
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0628 -0.3501 0.0152 0.3877 1.1399
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2558 0.5790 2.17 0.034 *
## FirstAuthorFemale1 0.0450 0.2077 0.22 0.829
## LastAuthorFemale1 -0.1855 0.1932 -0.96 0.341
## Year1997 -0.1604 0.5970 -0.27 0.789
## Year1998 -0.4877 0.6407 -0.76 0.449
## Year1999 -0.0915 0.5859 -0.16 0.876
## Year2000 -0.1248 0.5790 -0.22 0.830
## Year2002 0.4012 0.5657 0.71 0.481
## Year2003 -0.1126 0.5974 -0.19 0.851
## Year2004 -0.2738 0.6343 -0.43 0.667
## Year2005 0.0786 0.5789 0.14 0.892
## Year2006 0.2880 0.9074 0.32 0.752
```

```

## Year2007          -0.1323      0.5991   -0.22    0.826
## Year2008           0.1829      0.5899    0.31    0.757
## Year2009          -0.1099      0.5907   -0.19    0.853
## Year2010          -0.1280      0.5769   -0.22    0.825
## Year2011          -0.3827      0.6018   -0.64    0.527
## Year2012          -0.5419      0.6795   -0.80    0.428
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.577
## Multiple R-squared:  0.163, Adjusted R-squared:  -0.0564
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.676  0.868  0.952  0.917  0.984  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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 1.327e+14 1      1.152e+07
## Year              1.327e+14 15      2.956e+00

```

Residuals from first author



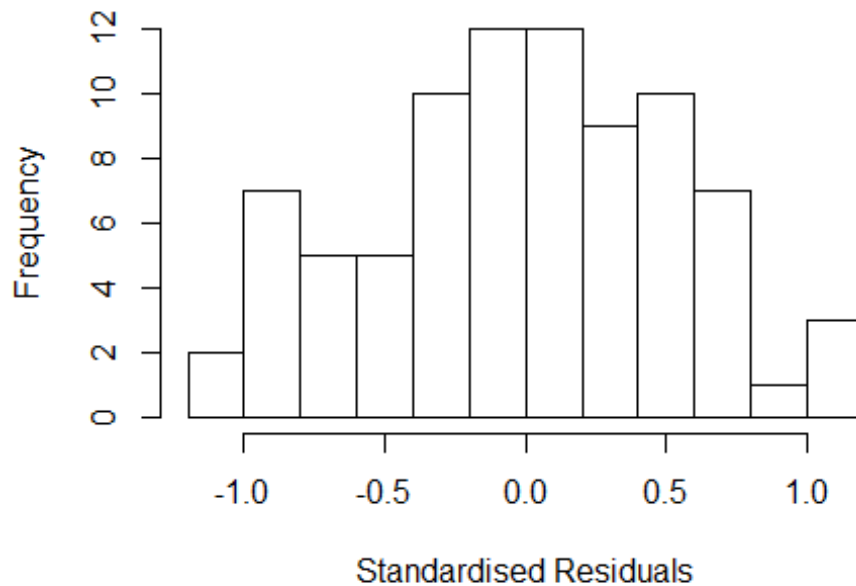
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0933 -0.3642 0.0351 0.3788 1.1501
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3054 0.7231 1.81 0.076 .
## FirstAuthorFemale1 -0.0974 0.1417 -0.69 0.494
## Year1997 -0.2640 0.7383 -0.36 0.722
## Year1998 -0.5372 0.7714 -0.70 0.489
## Year1999 -0.1412 0.7287 -0.19 0.847
## Year2000 -0.1744 0.7231 -0.24 0.810
## Year2002 0.4940 0.7090 0.70 0.488
## Year2003 -0.1905 0.7424 -0.26 0.798
## Year2004 -0.3275 0.7865 -0.42 0.678
## Year2005 0.0586 0.7359 0.08 0.937
## Year2006 0.2168 1.0329 0.21 0.834
## Year2007 -0.1822 0.7314 -0.25 0.804
```

```

## Year2008          0.1639      0.7363      0.22      0.825
## Year2009          -0.2116      0.7274     -0.29      0.772
## Year2010          -0.1511      0.7280     -0.21      0.836
## Year2011          -0.4425      0.7361     -0.60      0.550
## Year2012          -0.6176      0.7899     -0.78      0.437
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.579
## Multiple R-squared:  0.156, Adjusted R-squared:  -0.049
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 76 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.673  0.876  0.954  0.920  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      1.20e-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 2.109 1          1.452
## Year             2.109 15          1.025

```

Residuals from last author



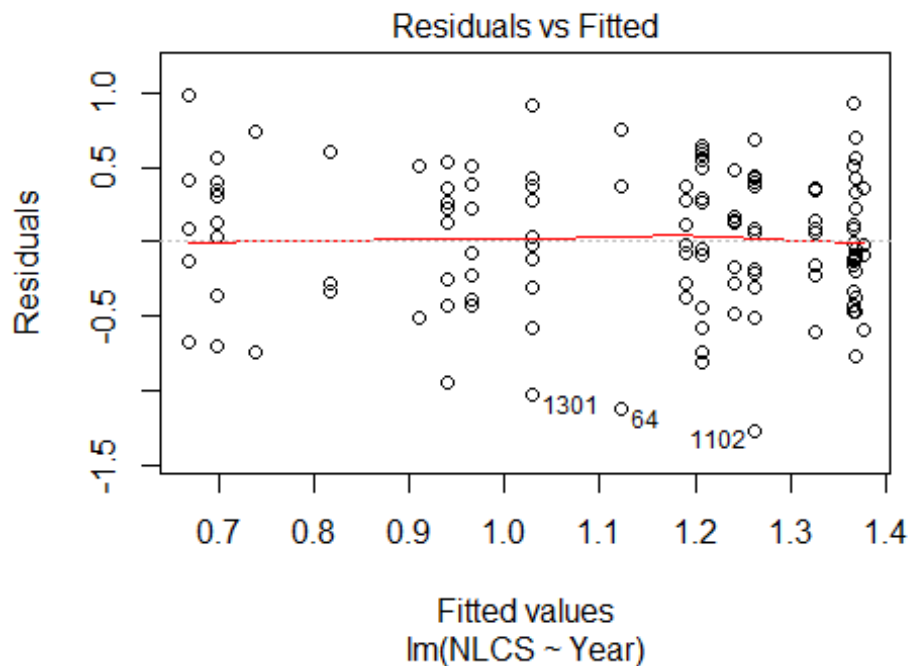
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.0788 -0.3507 0.0209 0.3877 1.1359
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2844 0.5775 2.22 0.03 *
## LastAuthorFemale1 -0.1529 0.1345 -1.14 0.26
## Year1997 -0.1939 0.5935 -0.33 0.74
## Year1998 -0.5158 0.6367 -0.81 0.42
## Year1999 -0.1202 0.5845 -0.21 0.84
## Year2000 -0.1534 0.5775 -0.27 0.79
## Year2002 0.4176 0.5775 0.72 0.47
## Year2003 -0.1349 0.6066 -0.22 0.82
## Year2004 -0.2877 0.6519 -0.44 0.66
## Year2005 0.0685 0.5926 0.12 0.91
## Year2006 0.2655 0.8913 0.30 0.77
## Year2007 -0.1527 0.5979 -0.26 0.80
```

```

## Year2008          0.1703      0.6116      0.28      0.78
## Year2009          -0.1398     0.5906     -0.24     0.81
## Year2010          -0.1415     0.5940     -0.24     0.81
## Year2011          -0.4073     0.6016     -0.68     0.50
## Year2012          -0.5608     0.6763     -0.83     0.41
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.591
## Multiple R-squared:  0.161, Adjusted R-squared:  -0.0422
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 73 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.692  0.876  0.956  0.921  0.988  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          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-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: 83"
## [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
##   56   69   71   66   50   44   76   50   74   63   67   65   59   74   54
## 2011 2012
##   53   77
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    7    7    3    2    5    6    7    9    8    9   11   12   14
## 2011 2012
##   14   10

```

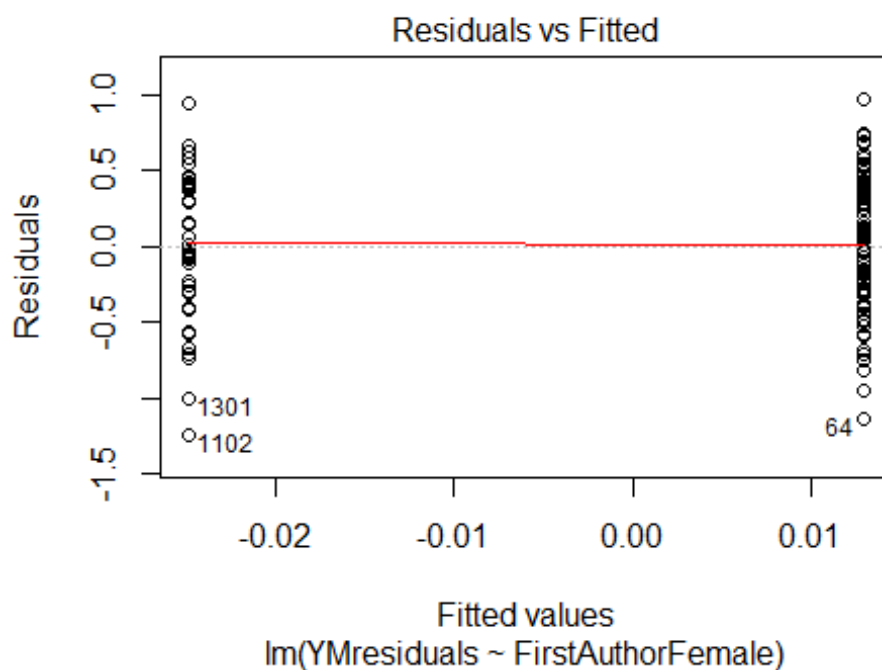
```
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    3    2    7    7    2    2    4    5    6    9    6    9    9   12   11
## 2011 2012
##   13    7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
##
## Bartlett test of homogeneity of variances
##
## data:  NLCS by Year
## Bartlett's K-squared = 15, df = 16, p-value = 0.5
```



```
##
## Bartlett test of homogeneity of variances
##
## data:  YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.99, df = 1, p-value = 0.3
## [1] "Female first author team size 2018 geometric mean: 3.6872946082583"
## [1] "Male first author team size 2018 geometric mean: 3.72161298232339"
## 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 = 76, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.14068083346529"
## [1] "Male last author team size 2018 geometric mean: 3.5893724679067"

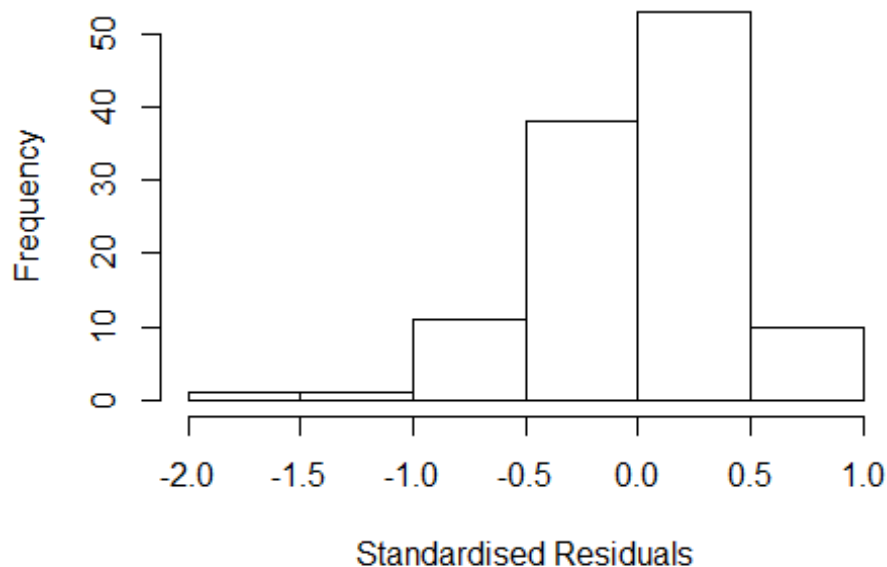
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 65, 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"
##
```

	GVI	F	Df	GVI ^{1/(2*Df)}
FirstAuthorFemale	2.339	1		1.529
LastAuthorFemale	1.687	1		1.299
UniqueAuthors	10.967	4		1.349
Year	13.241	16		1.084

Residuals from first and last author and team size



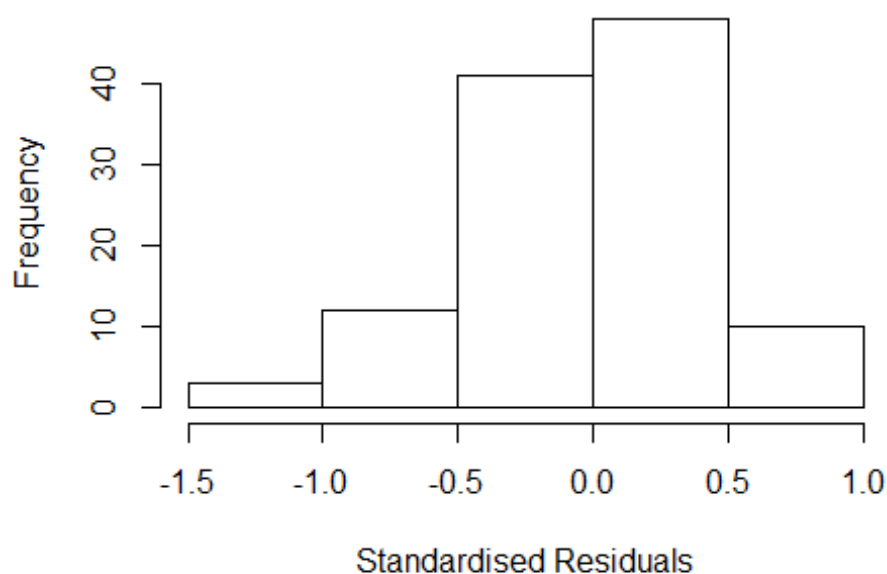
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
## UniqueAuthors +
## Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
## k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.8107 -0.3613 0.0367 0.2902 0.9151
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.5356 0.2202 6.97 4.8e-10 ***
## FirstAuthorFemale1 0.0646 0.1182 0.55 0.58626
## LastAuthorFemale1 -0.0777 0.1233 -0.63 0.53018
## UniqueAuthors2 0.2319 0.2131 1.09 0.27926
## UniqueAuthors3 0.3323 0.1897 1.75 0.08316 .
## UniqueAuthors4 0.1785 0.1881 0.95 0.34529
## UniqueAuthors5 0.2751 0.2105 1.31 0.19451
## Year1997 -0.7890 1.2388 -0.64 0.52577
## Year1998 -0.7957 0.2848 -2.79 0.00636 **
## Year1999 -0.4885 0.2338 -2.09 0.03949 *
```

```

## Year2000          -0.8414      0.3921    -2.15    0.03455 *
## Year2001          -0.6241      0.5502    -1.13    0.25964
## Year2002          -0.2522      0.3189    -0.79    0.43106
## Year2003          -1.0763      0.3879    -2.77    0.00670 **
## Year2004          -0.5508      0.2466    -2.23    0.02797 *
## Year2005          -0.8100      0.2605    -3.11    0.00250 **
## Year2006          -0.3987      0.2621    -1.52    0.13173
## Year2007          -1.0548      0.3032    -3.48    0.00078 ***
## Year2008          -0.4808      0.2872    -1.67    0.09757 .
## Year2009          -0.4202      0.2799    -1.50    0.13671
## Year2010          -0.3324      0.2799    -1.19    0.23806
## Year2011          -0.3824      0.2871    -1.33    0.18623
## Year2012          -0.6177      0.3103    -1.99    0.04955 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.447
## Multiple R-squared:  0.286, Adjusted R-squared:  0.114
## Convergence in 26 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 104 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0633 0.9010 0.9470 0.9070 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          8.77e-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 1.623 1          1.274
## LastAuthorFemale 1.401 1          1.184
## Year              1.861 16          1.020

```

Residuals from first and last author



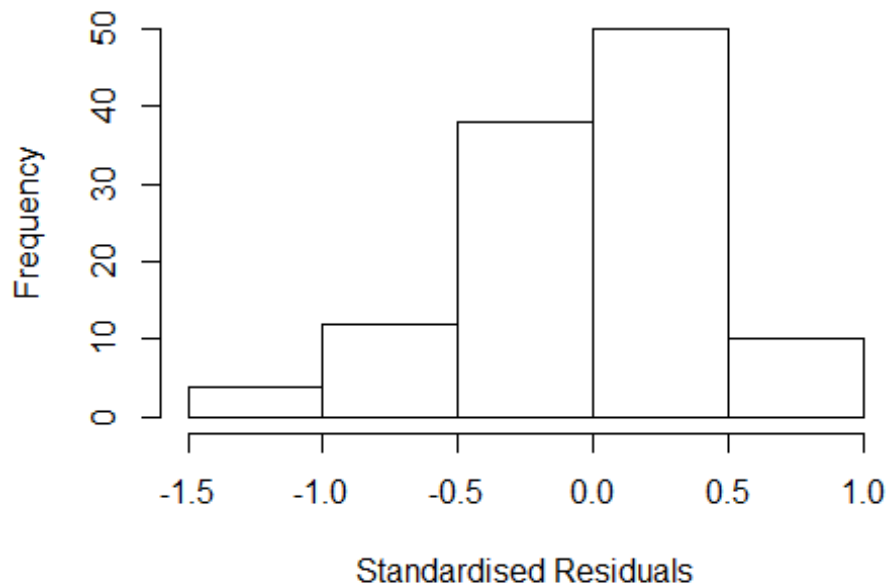
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.4587 -0.3067 0.0252 0.3224 0.9847
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4587 0.5718 2.55 0.012 *
## FirstAuthorFemale1 0.0755 0.1069 0.71 0.482
## LastAuthorFemale1 -0.0683 0.1217 -0.56 0.576
## Year1997 -0.7223 1.3362 -0.54 0.590
## Year1998 -0.5092 0.6056 -0.84 0.403
## Year1999 -0.2611 0.5844 -0.45 0.656
## Year2000 -0.5164 0.6473 -0.80 0.427
## Year2001 -0.5472 0.7183 -0.76 0.448
## Year2002 0.0474 0.5840 0.08 0.935
## Year2003 -0.7954 0.7257 -1.10 0.276
## Year2004 -0.3093 0.5847 -0.53 0.598
## Year2005 -0.4833 0.5907 -0.82 0.415
```

```

## Year2006          -0.0689      0.5830   -0.12    0.906
## Year2007          -0.7579      0.5897   -1.29    0.202
## Year2008          -0.1251      0.6078   -0.21    0.837
## Year2009          -0.1271      0.5835   -0.22    0.828
## Year2010          -0.0356      0.5807   -0.06    0.951
## Year2011          -0.0789      0.5965   -0.13    0.895
## Year2012          -0.3205      0.6153   -0.52    0.604
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.457
## Multiple R-squared:  0.24,   Adjusted R-squared:  0.0956
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 105 ones are summarized as
##   Min. 1st Qu.  Median      Mean 3rd Qu.    Max.
##  0.286  0.903  0.953  0.909  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      8.77e-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 1.401 1          1.184
## Year              1.401 16          1.011

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.4500 -0.3413 0.0404 0.3076 0.9840
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.4500 0.5776 2.51 0.014 *
## FirstAuthorFemale1 0.0796 0.1042 0.76 0.446
## Year1997 -0.7499 1.4793 -0.51 0.613
## Year1998 -0.5130 0.6048 -0.85 0.398
## Year1999 -0.2818 0.5840 -0.48 0.631
## Year2000 -0.5099 0.6803 -0.75 0.455
## Year2001 -0.5385 0.7586 -0.71 0.479
## Year2002 0.0551 0.5899 0.09 0.926
## Year2003 -0.7861 0.7260 -1.08 0.282
## Year2004 -0.3131 0.5857 -0.53 0.594
## Year2005 -0.5115 0.5878 -0.87 0.386
## Year2006 -0.0819 0.5834 -0.14 0.889
```

```

## Year2007          -0.7707      0.5905   -1.31    0.195
## Year2008          -0.1365      0.6034   -0.23    0.821
## Year2009          -0.1420      0.5803   -0.24    0.807
## Year2010          -0.0435      0.5831   -0.07    0.941
## Year2011          -0.0911      0.5963   -0.15    0.879
## Year2012          -0.3239      0.6182   -0.52    0.602
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.46
## Multiple R-squared:  0.238, Adjusted R-squared:  0.103
## Convergence in 27 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 105 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.299  0.896  0.953  0.909  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      8.77e-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 1.231 1      1.109
## 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 ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##      control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.4752 -0.3157  0.0204  0.3277  1.0113

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4752    0.5525   2.67  0.0089 **
## LastAuthorFemale1 -0.0719    0.1183  -0.61  0.5444
## Year1997          -0.6993    1.1142  -0.63  0.5318
## Year1998          -0.5019    0.5868  -0.86  0.3945
## Year1999          -0.2558    0.5656  -0.45  0.6521
## Year2000          -0.4952    0.6870  -0.72  0.4727
## Year2001          -0.5637    0.7046  -0.80  0.4256
## Year2002           0.0492    0.5627   0.09  0.9306
## Year2003          -0.8132    0.7148  -1.14  0.2581
## Year2004          -0.2992    0.5631  -0.53  0.5964
## Year2005          -0.4598    0.5648  -0.81  0.4177
## Year2006          -0.0843    0.5640  -0.15  0.8815
## Year2007          -0.7290    0.5650  -1.29  0.2000
## Year2008          -0.1256    0.5934  -0.21  0.8328
## Year2009          -0.1232    0.5665  -0.22  0.8282
## Year2010          -0.0252    0.5578  -0.05  0.9640
## Year2011          -0.0782    0.5749  -0.14  0.8921
## Year2012          -0.2979    0.5969  -0.50  0.6188
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.451
## Multiple R-squared:  0.236, Adjusted R-squared:  0.101
## Convergence in 24 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~ = 1. The remaining 106 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.263  0.897  0.951  0.907  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      8.77e-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: 114"
## [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 1999 2001 2002 2003 2006 2007 2009 2010 2011
##    4    3    1    1    4    3    1    1    1    1    1
##
## 1996 1997 1999 2001 2002 2003 2006 2007 2009 2010 2011
##    0    1    1    0    2    1    1    0    1    0    0
##
## 1996 1997 1999 2001 2002 2003 2006 2007 2009 2010 2011
##    0    1    1    0    2    0    1    0    1    0    0
## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 6"
## [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 2009 2010 2011
##    4    2   10    7    1    1    2    1    3    3    1    2    2    3    1
## 2012
##    5
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011
##    1    1    1    1    0    0    1    0    1    0    0    1    0    0    0
## 2012
##    0
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011
##    1    1    0    0    0    0    0    0    1    0    0    1    0    0    0
## 2012
##    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"

```



```

## [1] "Female first author team size 2018 geometric mean: 6"
## [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: 4"
## [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
##    8   12   13   12   13    3    6    6    5    9    2   11   16    8    4
## 2011 2012
##   16    8
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    2    1    1    0    0    2    2    1    0    2    4    4    1    0
## 2011 2012
##    7    1
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    1    1    0    0    1    0    1    0    2    4    4    1    0
## 2011 2012
##    6    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.64621617328617"
## [1] "Male first author team size 2018 geometric mean: 2.82842712474619"

## 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 = 6, p-value = 0.1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 4.28225473667665"
## [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: 24"
## [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 2007 2009 2010 2011 2012
##    4    2    9    1    3    2    2    1    3    5    3    1    5    2    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2007 2009 2010 2011 2012
##    2    1    2    0    0    0    1    0    1    1    1    0    2    1    0
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2007 2009 2010 2011 2012
##    2    1    2    0    0    0    0    0    1    1    1    0    2    1    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 4.24264068711928"
## [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: 11"
## [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
##   15    3   17    5   10   15   15    7   14    7   12   14   16   15   15
## 2011 2012
##   26   20
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    2    1    4    1    0    3    4    3    4    2    4    5   11    4    9
## 2011 2012
##   13   10
##

```

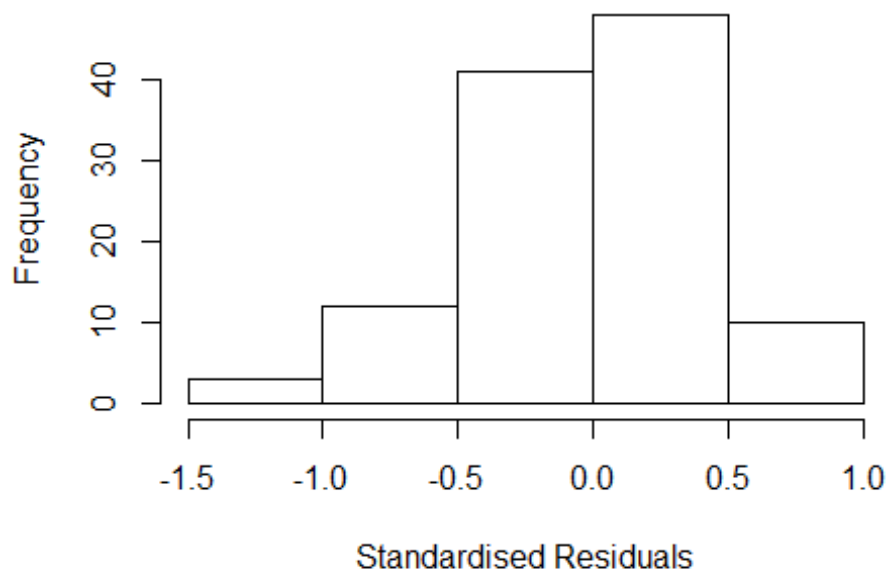
```
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##      2      1      4      1      0      1      2      3      4      2      3      4      5      4      9
## 2011 2012
##     11      7
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 6"
## [1] "Male first author team size 2018 geometric mean: 5.14352079675504"

## 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 = 0.3
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 5"
## [1] "Male last author team size 2018 geometric mean: 5.4400868459948"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```

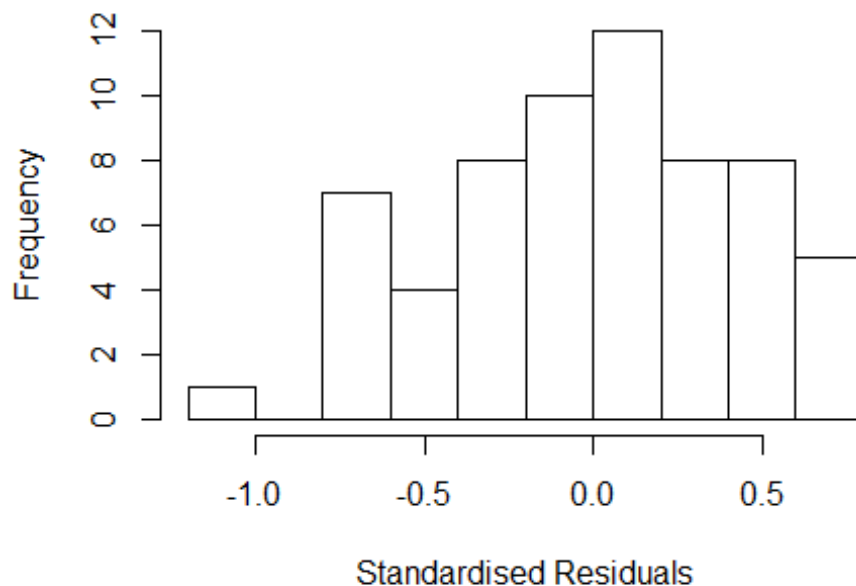
Residuals from last author



```
##
## Wilcoxon rank sum test with continuity correction
```

```
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1.5, 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 2.319e+13 1      4.816e+06
## LastAuthorFemale 2.306e+13 1      4.802e+06
## UniqueAuthors    1.523e+15 4      7.904e+01
## Year              3.146e+28 15     8.911e+00
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.0352 -0.2515 0.0137 0.2659 0.7719
```

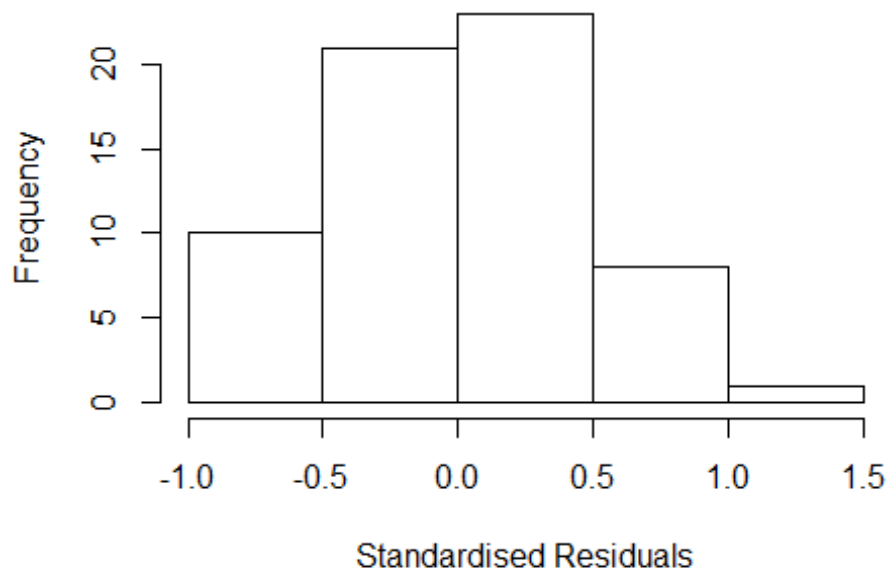
```

##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.09662    0.91869   1.19  0.23946
## FirstAuthorFemale1 -0.00771    0.15616  -0.05  0.96085
## LastAuthorFemale1 -0.08703    0.15570  -0.56  0.57924
## UniqueAuthors2    0.11475    0.20336   0.56  0.57564
## UniqueAuthors3    0.70537    0.19001   3.71  0.00061 ***
## UniqueAuthors4    0.55520    0.27597   2.01  0.05085 .
## UniqueAuthors5    0.59585    0.31382   1.90  0.06465 .
## Year1997        -0.71899    0.92488  -0.78  0.44139
## Year1998        -0.28121    0.95872  -0.29  0.77076
## Year1999        -1.09662    0.91869  -1.19  0.23946
## Year2001        -0.71488    0.93423  -0.77  0.44853
## Year2002        -0.61794    0.94648  -0.65  0.51748
## Year2003        -0.43504    0.94033  -0.46  0.64606
## Year2004        -0.23113    0.92779  -0.25  0.80451
## Year2005        -0.13137    0.92572  -0.14  0.88785
## Year2006        -0.61020    0.93144  -0.66  0.51605
## Year2007        -0.42462    0.96177  -0.44  0.66117
## Year2008        -0.71869    0.92610  -0.78  0.44218
## Year2009        -0.91777    0.93938  -0.98  0.33430
## Year2010        -0.61658    0.93767  -0.66  0.51449
## Year2011        -0.28278    0.93267  -0.30  0.76327
## Year2012        -0.65034    0.94049  -0.69  0.49316
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.491
## Multiple R-squared:  0.383, Adjusted R-squared:  0.0677
## Convergence in 21 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 57 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.636  0.903  0.969  0.931  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           1.59e-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 -4.821e+14  1      NaN
## LastAuthorFemale  -3.810e+14  1      NaN
## Year              -6.921e+14 15      NaN
```

Residuals from first and last author



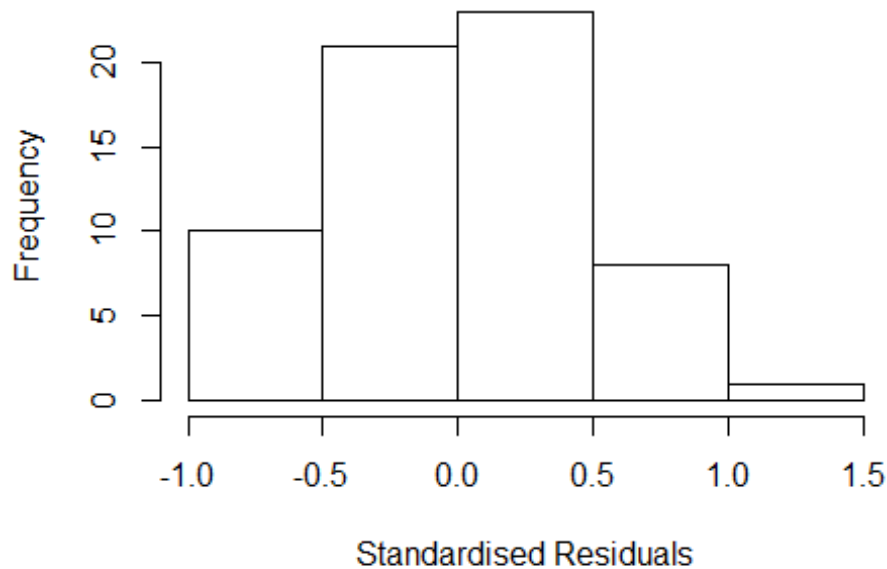
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.9666 -0.3370  0.0178  0.3073  1.0857
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.1540     0.6288   1.84    0.073 .
## FirstAuthorFemale1  0.0660     0.1636   0.40    0.688
## LastAuthorFemale1  0.0335     0.1626   0.21    0.838
```

```

## Year1997          -0.0710      0.6288   -0.11    0.911
## Year1998          -0.2870      0.7321   -0.39    0.697
## Year1999          -1.1540      0.6288   -1.84    0.073 .
## Year2001          -0.9665      0.6455   -1.50    0.141
## Year2002          -0.4198      0.6329   -0.66    0.511
## Year2003          -0.3533      0.6893   -0.51    0.611
## Year2004           0.0602      0.6544    0.09    0.927
## Year2005           0.0520      0.6745    0.08    0.939
## Year2006          -0.4753      0.7023   -0.68    0.502
## Year2007          -0.1916      0.6955   -0.28    0.784
## Year2008          -0.5497      0.6921   -0.79    0.431
## Year2009          -0.5354      0.6441   -0.83    0.410
## Year2010          -0.2989      0.6643   -0.45    0.655
## Year2011          -0.1137      0.6509   -0.17    0.862
## Year2012          -0.3366      0.6712   -0.50    0.618
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.587
## Multiple R-squared:  0.175, Adjusted R-squared:  -0.137
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.713  0.896  0.966  0.930  0.985  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      1.59e-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 3.169e+13 1      5.630e+06
## Year              3.169e+13 15      2.819e+00

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -0.9713 -0.3285 0.0183 0.3068 1.0775
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1540 0.6448 1.79 0.08 .
## FirstAuthorFemale1 0.0838 0.1405 0.60 0.55
## Year1997 -0.0710 0.6448 -0.11 0.91
## Year1998 -0.2813 0.7463 -0.38 0.71
## Year1999 -1.1540 0.6448 -1.79 0.08 .
## Year2001 -0.9508 0.6571 -1.45 0.15
## Year2002 -0.4119 0.6476 -0.64 0.53
## Year2003 -0.3426 0.7051 -0.49 0.63
## Year2004 0.0597 0.6701 0.09 0.93
## Year2005 0.0431 0.6856 0.06 0.95
## Year2006 -0.4818 0.7154 -0.67 0.50
## Year2007 -0.1827 0.7109 -0.26 0.80
```



```

## Year2008          -0.5531      0.7069   -0.78      0.44
## Year2009          -0.5400      0.6600   -0.82      0.42
## Year2010          -0.2876      0.6757   -0.43      0.67
## Year2011          -0.1055      0.6648   -0.16      0.87
## Year2012          -0.3322      0.6849   -0.49      0.63
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.568
## Multiple R-squared:  0.176, Adjusted R-squared:  -0.111
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 54 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.699  0.890  0.960  0.925  0.983  0.997
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+00          5.00e-01      4.69e+00      1.00e-07
##      rel.tol          solve.tol      eps.outlier      eps.x
##      1.00e-07          1.00e-07      1.59e-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 -7.594e+13  1      NaN
## Year              -7.594e+13 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
## -9.73e-01 -3.43e-01 -1.11e-16  3.20e-01  1.07e+00
##

```

```

## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)    1.1540    0.6326   1.82   0.075 .
## LastAuthorFemale1 0.0677    0.1368   0.50   0.623
## Year1997       -0.0710    0.6326  -0.11   0.911
## Year1998       -0.2581    0.7482  -0.35   0.732
## Year1999       -1.1540    0.6326  -1.82   0.075 .
## Year2001       -0.9347    0.6450  -1.45   0.154
## Year2002       -0.4039    0.6354  -0.64   0.528
## Year2003       -0.3337    0.6927  -0.48   0.632
## Year2004        0.0859    0.6571   0.13   0.897
## Year2005        0.0850    0.6843   0.12   0.902
## Year2006       -0.4602    0.7206  -0.64   0.526
## Year2007       -0.1810    0.7006  -0.26   0.797
## Year2008       -0.5261    0.6857  -0.77   0.447
## Year2009       -0.4950    0.6411  -0.77   0.444
## Year2010       -0.2988    0.6654  -0.45   0.656
## Year2011       -0.1024    0.6542  -0.16   0.876
## Year2012       -0.3258    0.6735  -0.48   0.631
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.582
## Multiple R-squared:  0.173, Adjusted R-squared:  -0.115
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 55 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.714  0.894  0.965  0.930  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      1.59e-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: 63"
## [1] ""
## [1] ""
## [1] "#####"

```

```

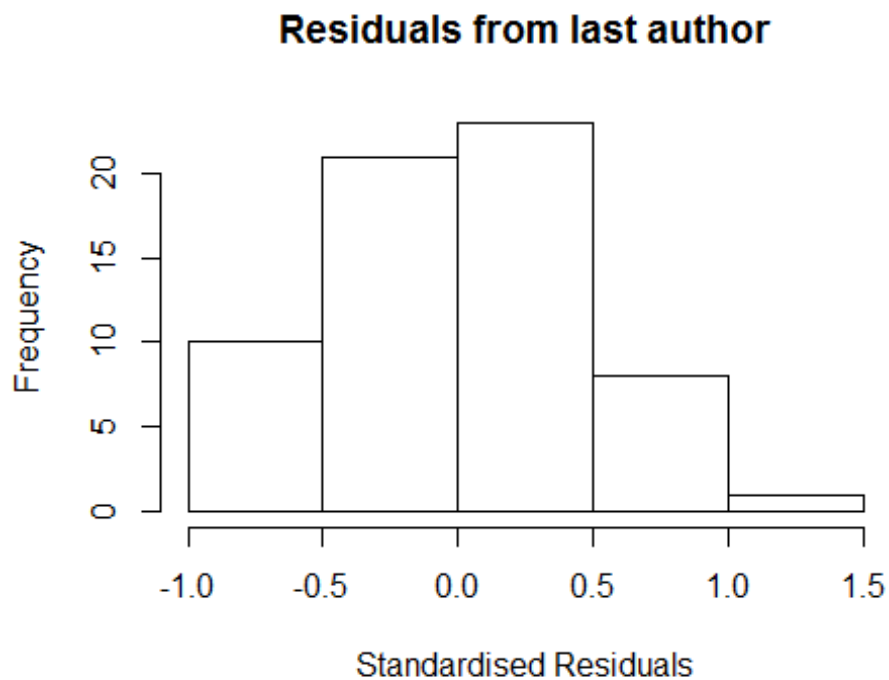
## [1] "Analysis of AJSC 3501"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1996 1998 2010
##    1    3    1
##
## 1996 1998 2010
##    0    0    1
##
## 1996 1998 2010
##    0    0    1
## [1] "Heteroscedasticity checks, 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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 1"
## [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
##    2    2    3    1    2    3    5    3    5    1    3    2    3    7    9
## 2011 2012
##    6    4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    2    0    0    3    3    2    3    0    3    1    1    4    5
## 2011 2012
##    5    3
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##    1    1    2    0    0    1    1    2    3    0    3    0    0    4    4
## 2011 2012
##    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: 3.46410161513775"

```

```
## [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"

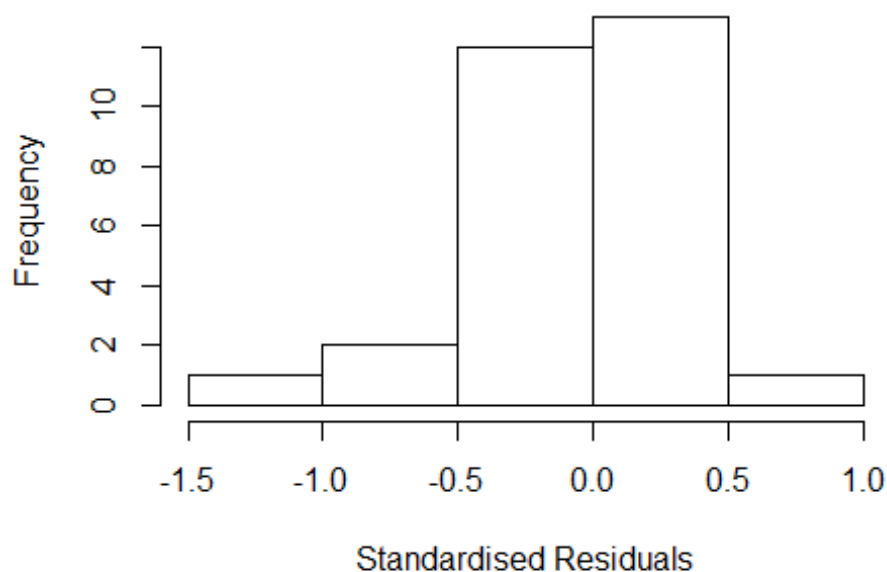
## Warning in lf.cov(init, x = x): .vcov.avar1: negative diag(<vcov>) fixed
## up; consider 'cov=".vcov.w."' instead

## 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
## LastAuthorFemale  NaN 1          NaN
## Year              NaN 11         NaN
```

Residuals from first and last author



```
## [1] "List of 0 outliers with 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
## -1.11e+00 -2.39e-01 1.50e-15 3.04e-01 5.80e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.4870 0.0000 Inf < 2e-16 ***
## FirstAuthorFemale1 -0.3155 0.2229 -1.42 0.1774
## LastAuthorFemale1 0.2408 0.2301 1.05 0.3120
## Year1997 0.5960 0.0000 Inf < 2e-16 ***
## Year1998 -0.1745 0.2397 -0.73 0.4779
## Year2001 0.5935 0.2229 2.66 0.0178 *
## Year2002 0.3818 0.2232 1.71 0.1078
## Year2003 0.4890 0.5723 0.85 0.4063
## Year2004 0.6140 0.2050 2.99 0.0091 **
## Year2006 0.6279 0.5625 1.12 0.2819
## Year2009 0.4979 0.3063 1.63 0.1249
## Year2010 1.2100 0.0753 16.07 7.3e-11 ***
```

```

## Year2011          0.9646      0.0722    13.37  9.8e-10 ***
## Year2012          1.0005      0.2525     3.96  0.0013 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.481
## Multiple R-squared:  0.523, Adjusted R-squared:  0.11
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 22 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.571  0.923  0.961  0.932  0.978  0.993
## Algorithmic parameters:
##      tuning.chi          bb      tuning.psi      refine.tol
##      1.55e+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-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 7.093e+14 1      2.663e+07
## Year              7.093e+14 11      4.732e+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 + Year, data =
AllScopusDataOlderFirstLastGendered,
##   control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min      1Q  Median      3Q      Max
## -1.27e+00 -1.67e-01  3.33e-16  1.71e-01  5.83e-01
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    4.87e-01   2.19e-08  2.22e+07 < 2e-16 ***

```

```

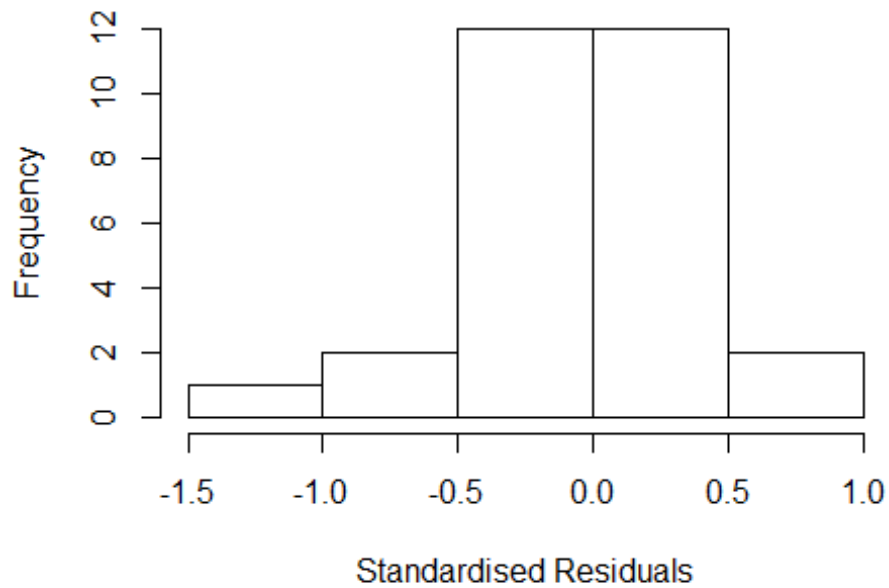
## FirstAuthorFemale1 -1.95e-01 1.12e-01 -1.75e+00 0.10008
## Year1997 5.96e-01 2.19e-08 2.72e+07 < 2e-16 ***
## Year1998 -1.74e-01 2.50e-01 -7.00e-01 0.49599
## Year2001 4.73e-01 1.12e-01 4.23e+00 0.00064 ***
## Year2002 5.02e-01 1.12e-01 4.49e+00 0.00037 ***
## Year2003 4.89e-01 7.29e-01 6.70e-01 0.51174
## Year2004 6.19e-01 2.21e-01 2.80e+00 0.01272 *
## Year2006 7.79e-01 6.61e-01 1.18e+00 0.25585
## Year2009 5.70e-01 4.32e-01 1.32e+00 0.20568
## Year2010 1.27e+00 3.57e-02 3.56e+01 < 2e-16 ***
## Year2011 9.65e-01 7.25e-02 1.33e+01 4.6e-10 ***
## Year2012 1.00e+00 2.65e-01 3.77e+00 0.00166 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.394
## Multiple R-squared: 0.555, Adjusted R-squared: 0.221
## Convergence in 18 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 23 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.282 0.896 0.944 0.902 0.990 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 3.45e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

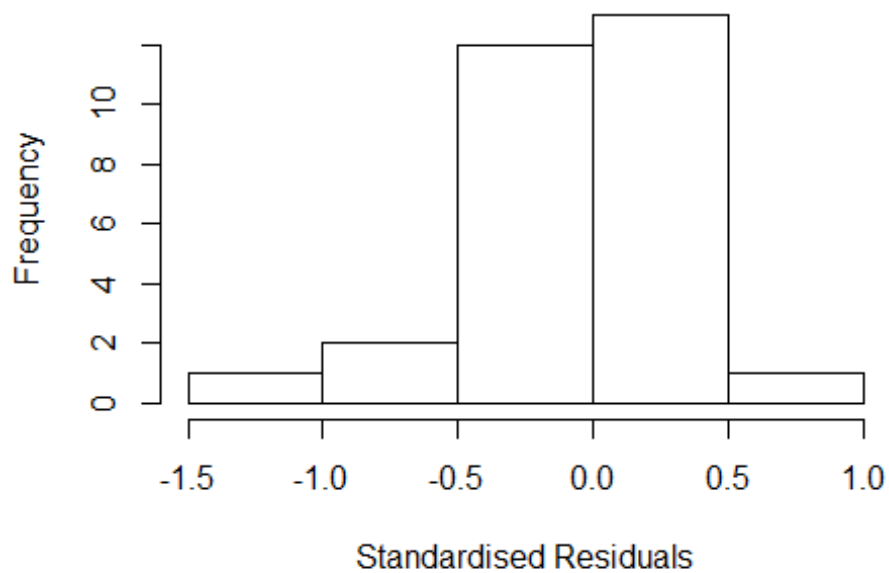
```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale  NaN  1          NaN
## Year              NaN 11          NaN
```

Residuals from last author




```

## [1] "List of 0 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
## -1.16e+00 -2.72e-01 1.67e-16 2.77e-01 5.80e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 4.87e-01 1.06e-08 4.62e+07 < 2e-16 ***
## LastAuthorFemale1 1.53e-01 2.49e-01 6.10e-01 0.5479
## Year1997 5.96e-01 8.61e-09 6.92e+07 < 2e-16 ***
## Year1998 -1.74e-01 2.43e-01 -7.20e-01 0.4836
## Year2001 2.78e-01 0.00e+00 Inf < 2e-16 ***
## Year2002 1.54e-01 2.49e-01 6.20e-01 0.5435
## Year2003 4.89e-01 6.17e-01 7.90e-01 0.4394
## Year2004 6.16e-01 2.11e-01 2.92e+00 0.0100 *
## Year2006 6.73e-01 6.56e-01 1.03e+00 0.3205
## Year2009 5.25e-01 3.43e-01 1.53e+00 0.1450
## Year2010 1.23e+00 7.54e-02 1.63e+01 2.1e-11 ***
## Year2011 8.59e-01 8.68e-02 9.89e+00 3.2e-08 ***
## Year2012 1.00e+00 2.57e-01 3.90e+00 0.0013 **
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.445
## Multiple R-squared: 0.522, Adjusted R-squared: 0.163
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~ = 1. The remaining 22 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.478 0.907 0.956 0.919 0.972 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 3.45e-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: 29"
## [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 2002 2006 2007 2008 2011 2012
##    1    1    1    2    1    1    1    1    1
##
## 1996 1997 1998 2002 2006 2007 2008 2011 2012
##    0    0    0    2    0    1    0    1    0
##
## 1996 1997 1998 2002 2006 2007 2008 2011 2012
##    0    0    0    1    0    1    0    1    0
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5"
## [1] "Male first author team size 2018 geometric mean: 6"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 5.47722557505166"
## [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: 3"
## [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]"
##

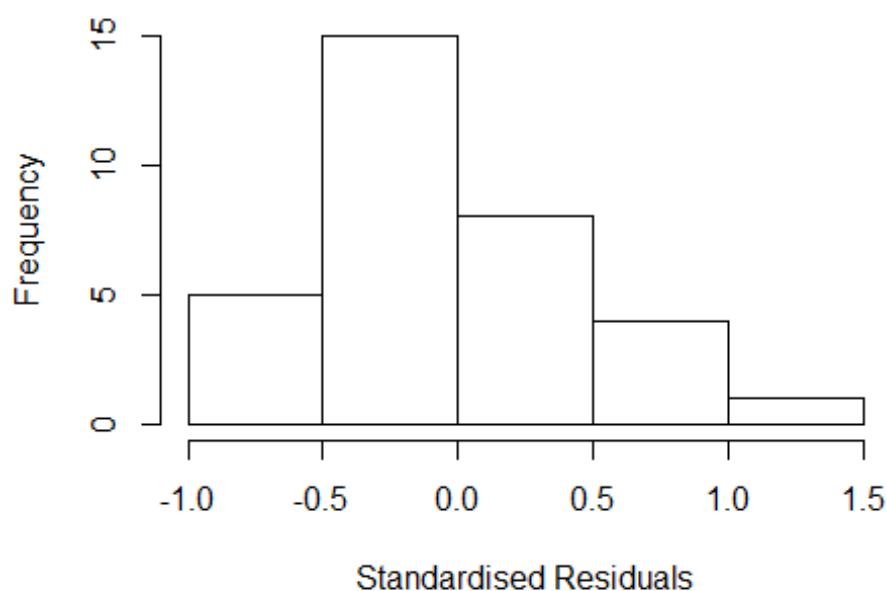
```

```

## 1996 1997 1998 1999 2002 2010
##    1    2    2    1    1    1
##
## 1996 1997 1998 1999 2002 2010
##    0    0    0    0    0    1
##
## 1996 1997 1998 1999 2002 2010
##    0    0    0    0    0    1
## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 1"
## [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]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011 2012
##    6    2    6    3    6    3    2    4    7    4    3    1    2    2    2
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011 2012
##    4    1    5    2    3    2    1    3    5    4    3    1    2    1    2
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011 2012
##    4    1    5    2    3    1    0    2    4    4    1    1    2    1    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"
##
##                               GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 1.587e+14  1      1.260e+07
## LastAuthorFemale  7.559e+14  1      2.749e+07
## Year              8.918e+28 13      1.299e+01

```

Residuals from first and last author



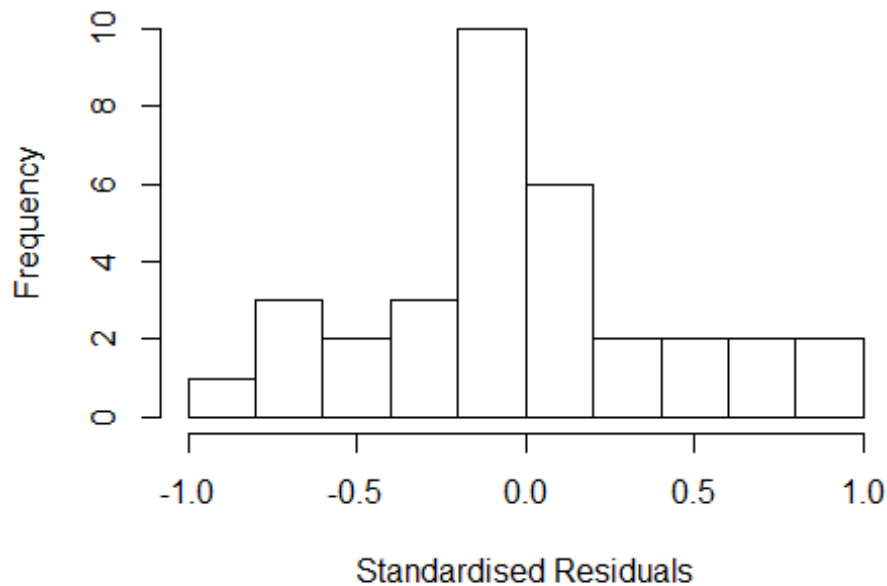
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -8.76e-01 -2.04e-01 5.55e-17 2.27e-01 1.09e+00
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2395 0.3616 3.43 0.0032 **
## FirstAuthorFemale1 0.2638 0.2007 1.31 0.2061
## LastAuthorFemale1 -0.2491 0.2045 -1.22 0.2398
## Year1998 0.0935 0.3804 0.25 0.8087
## Year1999 0.1005 0.4397 0.23 0.8219
## Year2000 0.3236 0.4045 0.80 0.4347
## Year2001 0.1478 0.3935 0.38 0.7119
## Year2002 0.0626 0.3931 0.16 0.8752
## Year2004 0.1642 0.5209 0.32 0.7564
## Year2005 -0.6886 0.3867 -1.78 0.0929 .
## Year2006 -0.3785 0.5430 -0.70 0.4952
## Year2007 -0.1665 0.3616 -0.46 0.6510
```

```

## Year2009          -0.3413      0.4449   -0.77    0.4536
## Year2010          0.2855      0.6553    0.44    0.6686
## Year2011          0.1977      0.4449    0.44    0.6623
## Year2012          0.8731      0.7869    1.11    0.2827
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.637
## Multiple R-squared:  0.364, Adjusted R-squared:  -0.198
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.751  0.909  0.984  0.945  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      3.03e-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 3.36e+15 1      5.797e+07
## Year              3.36e+15 13      3.955e+00

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -9.57e-01 -2.58e-01 8.33e-17 1.90e-01 9.98e-01
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.1640 0.3254 3.58 0.0022 **
## FirstAuthorFemale1 0.1817 0.2567 0.71 0.4880
## Year1998 -0.0800 0.3254 -0.25 0.8085
## Year1999 0.0667 0.4016 0.17 0.8699
## Year2000 0.1911 0.3427 0.56 0.5840
## Year2001 0.1110 0.3748 0.30 0.7704
## Year2002 0.2202 0.3868 0.57 0.5762
## Year2004 0.1972 0.6205 0.32 0.7543
## Year2005 -0.5729 0.3523 -1.63 0.1213
## Year2006 -0.3888 0.5120 -0.76 0.4575
## Year2007 -0.0910 0.3254 -0.28 0.7828
## Year2009 -0.4328 0.3868 -1.12 0.2779
```

```

## Year2010          0.3610      0.6519      0.55      0.5866
## Year2011          0.1062      0.3868      0.27      0.7867
## Year2012          0.8651      0.9270      0.93      0.3631
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.611
## Multiple R-squared:  0.341, Adjusted R-squared:  -0.172
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 25 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.772  0.902  0.977  0.934  0.993  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      3.03e-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 -5.785e+13  1      NaN
## Year              -5.785e+13 13      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
## -9.05e-01 -2.33e-01 -3.82e-16  2.01e-01  1.20e+00
##
## Coefficients:
##      Estimate Std. Error t value Pr(>|t|)

```

```

## (Intercept)          1.3092      0.3932      3.33      0.0037 **
## LastAuthorFemale1    -0.1641      0.2571     -0.64      0.5314
## Year1998              -0.0611      0.4850     -0.13      0.9011
## Year1999              0.0600      0.4666      0.13      0.8991
## Year2000              0.3009      0.5062      0.59      0.5597
## Year2001              0.1973      0.4438      0.44      0.6619
## Year2002              0.2568      0.3932      0.65      0.5220
## Year2004              0.3158      0.6057      0.52      0.6085
## Year2005             -0.6294      0.4101     -1.53      0.1423
## Year2006             -0.2406      0.5463     -0.44      0.6648
## Year2007             -0.2362      0.3932     -0.60      0.5554
## Year2009             -0.2321      0.4850     -0.48      0.6379
## Year2010              0.2158      0.6817      0.32      0.7553
## Year2011              0.3069      0.4850      0.63      0.5349
## Year2012              0.8928      0.6728      1.33      0.2011
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.614
## Multiple R-squared:  0.348, Adjusted R-squared:  -0.159
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 6 weights are ~= 1. The remaining 27 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.682  0.908  0.986  0.938  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          3.03e-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: 33"
## [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]"

```



```

##
## 1998 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011 2012
##    1    1    2    2    2    1    1    2    1    3    5    4    7
##
## 1998 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011 2012
##    1    1    1    0    1    1    1    2    0    3    5    4    5
##
## 1998 2000 2001 2002 2003 2004 2005 2006 2007 2009 2010 2011 2012
##    1    1    1    0    1    1    1    2    0    3    5    3    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.44948974278318"
## [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: 24"
## [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]"
##
## 2000 2005 2006 2008 2011 2012
##    1    2    2    1    1    1
##
## 2000 2005 2006 2008 2011 2012
##    1    2    2    1    1    1
##
## 2000 2005 2006 2008 2011 2012
##    1    2    2    1    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.47722557505166"
## [1] "Male first author team size 2018 geometric mean: 6"

## 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 = 1
## alternative hypothesis: true location shift is not equal to 0

```

```

##
## [1] "Female last author team size 2018 geometric mean: 6"
## [1] "Male last author team size 2018 geometric mean: 5"

## 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 = 2, 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"
## [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: 8"
## [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]"
##
## 2002 2003 2005 2006 2008 2009 2011
## 1 1 1 2 2 3 1
##
## 2002 2003 2005 2006 2008 2009 2011
## 1 1 1 2 2 2 1
##
## 2002 2003 2005 2006 2008 2009 2011
## 0 1 1 2 2 0 1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [1] "Male first author team size 2018 geometric mean: 1"
##
## Wilcoxon rank sum test
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 1, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: NaN"
## [1] "Male last author team size 2018 geometric mean: 1.73205080756888"

```

```

## [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: 7"
## [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]"
##
## 2000 2001 2002 2003 2004 2008 2009 2010 2011 2012
##    1    2    2    1    1    2    1    2    2    1
##
## 2000 2001 2002 2003 2004 2008 2009 2010 2011 2012
##    1    1    0    0    1    2    1    2    1    1
##
## 2000 2001 2002 2003 2004 2008 2009 2010 2011 2012
##    1    1    0    0    1    2    1    2    1    1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [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

## [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: 10"
## [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]"
##
## 2000 2001 2003 2004 2005 2006 2008 2009 2011 2012
##    1    1    1    3    2    3    2    3    3    1
##
## 2000 2001 2003 2004 2005 2006 2008 2009 2011 2012

```

```

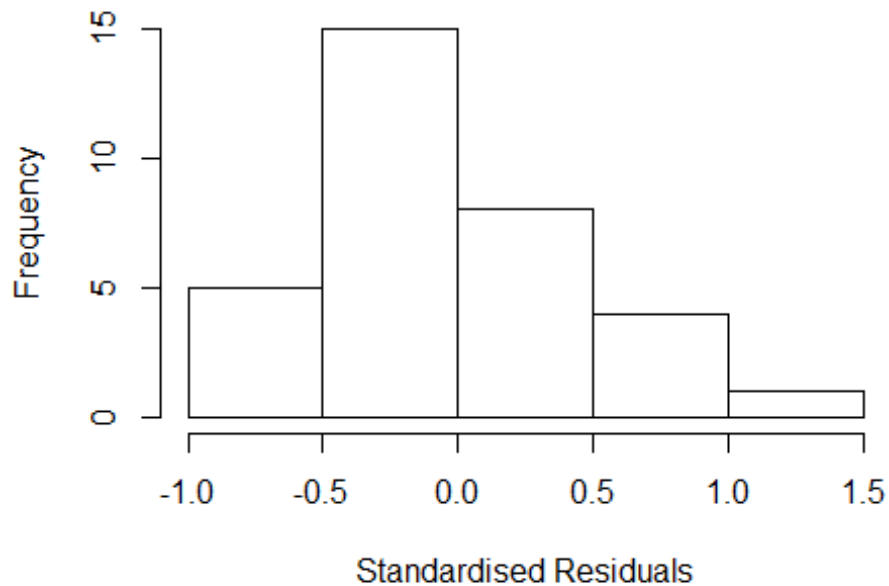
##      0      0      1      2      2      3      2      1      2      1
##
## 2000 2001 2003 2004 2005 2006 2008 2009 2011 2012
##      0      0      1      1      1      3      2      1      1      1
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: NaN"
## [1] "Male first author team size 2018 geometric mean: 3"
## [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: 11"
## [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]"
##
## 1999 2000 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##      1     15      3      3      2      2      2      2      2      1      1      2
##
## 1999 2000 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##      0      7      1      2      1      1      2      0      1      1      0      2
##
## 1999 2000 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##      0      6      0      2      1      1      2      0      1      0      0      0
## [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: 4"
## [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

```

Residuals from last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1           NaN
## LastAuthorFemale  NaN 1           NaN
## Year              NaN 5           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:
##      2      6      7     12     13     15     23     25     26     28
## 0.000 0.000 0.000 0.000 0.000 0.000 0.000 -0.503 0.000 0.000
##     30     31     37
## 0.000 -1.181 0.000
##
## 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
```

```

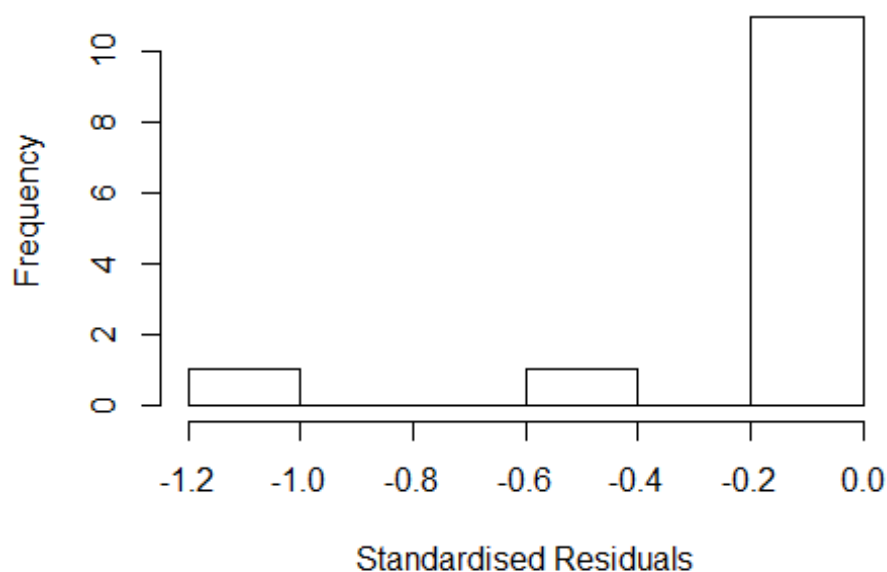
## Year2004          1.01      0.00      NA      NA
## Year2005          1.92      0.00      NA      NA
## Year2006          1.29      0.00      NA      NA
## Year2007          1.71      0.00      NA      NA
## Year2009          1.85      0.00      NA      NA
##
## Robustness weights:
## 2 observations c(8,12) are outliers with |weight| = 0 ( < 0.0077);
## 11 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          7.69e-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"

## 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

```

Residuals from first and last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## Year              NaN 5             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
## -0.503  0.000  0.000  0.000  1.181
##
## 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
## Year2004          1.01      0.00      NA      NA
## Year2005          1.92      0.00      NA      NA
## Year2006          1.29      0.00      NA      NA
## Year2007          0.53      0.00      NA      NA
```

```

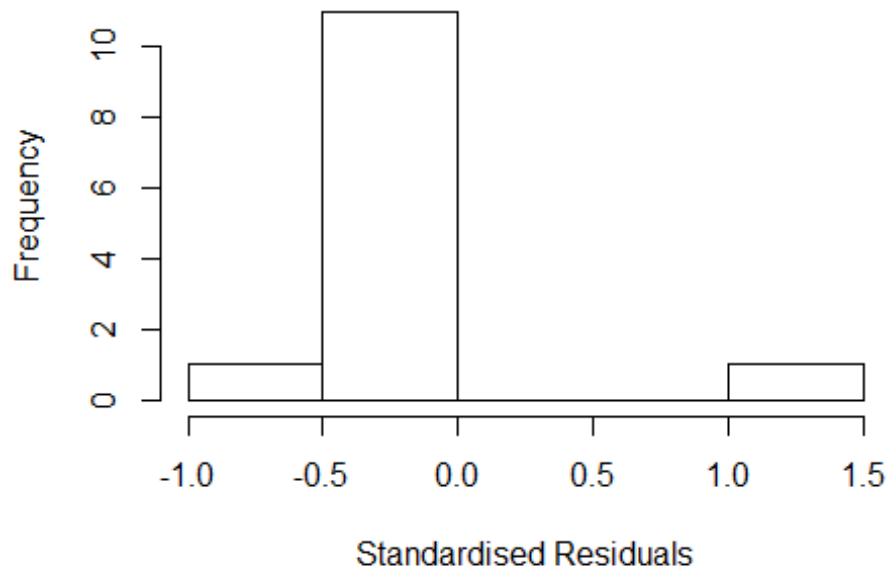
## Year2009          1.85          0.00          NA          NA
##
## Robustness weights:
## 2 observations c(8,11) are outliers with |weight| = 0 ( < 0.0077);
## 11 weights are ~= 1.
## Algorithmic parameters:
##      tuning.chi          bb          tuning.psi          refine.tol
##      1.55e+00          5.00e-01          4.69e+00          1.00e-07
##      rel.tol          solve.tol          eps.outlier          eps.x
##      1.00e-07          1.00e-07          7.69e-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 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

```


Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale NaN 1          NaN
## Year            NaN 5          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
## -0.503  0.000  0.000  0.000  1.181
##
## 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
## Year2004              1.01         0.00    NA      NA
## Year2005              1.92         0.00    NA      NA
## Year2006              1.29         0.00    NA      NA
## Year2007              0.53         0.00    NA      NA
```

```

## Year2009          1.85          0.00          NA          NA
##
## Robustness weights:
## 2 observations c(8,11) are outliers with |weight| = 0 ( < 0.0077);
## 11 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          7.69e-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: 13"
## [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 2000 2001 2003 2005 2007 2008 2009 2010 2011 2012
##    1    2    1    1    2    1    2    3    3    2    2    2    4
##
## 1996 1997 1998 2000 2001 2003 2005 2007 2008 2009 2010 2011 2012
##    1    1    1    1    2    1    1    2    3    1    2    2    4
##
## 1996 1997 1998 2000 2001 2003 2005 2007 2008 2009 2010 2011 2012
##    1    1    1    1    2    1    1    2    3    0    2    2    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3"
## [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): find_scale() did not
## converge in 'maxit.scale' (= 200) iterations

## [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: 20"
## [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]"
##
## 1997 1998 1999 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    4    2    3    2    2    3    2    3    5    3    4    1    1    6
##
## 1997 1998 1999 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    2    2    2    1    2    2    3    4    2    4    1    1    5
##
## 1997 1998 1999 2001 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    2    2    2    2    1    1    2    2    4    2    4    1    1    5
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 5.24148278841779"
## [1] "Male first author team size 2018 geometric mean: 2.91295063024394"

## 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 = 12, p-value = 0.07
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 4.24264068711928"
## [1] "Male last author team size 2018 geometric mean: 3.1748021039364"

## 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 = 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"

```

```

## [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: 31"
## [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]"
##
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    1    1    1    1    5    4   12    4    4    9   10    5
##
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    1    1    4    4   11    4    4    8   10    4
##
## 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012
##    0    1    1    1    3    4    9    4    3    8    9    3
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 3.63424118566428"
## [1] "Male first author team size 2018 geometric mean: 3.30192724889463"

## 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 = 9.5, p-value = 1
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 3.88832259447933"
## [1] "Male last author team size 2018 geometric mean: 2.88449914061482"

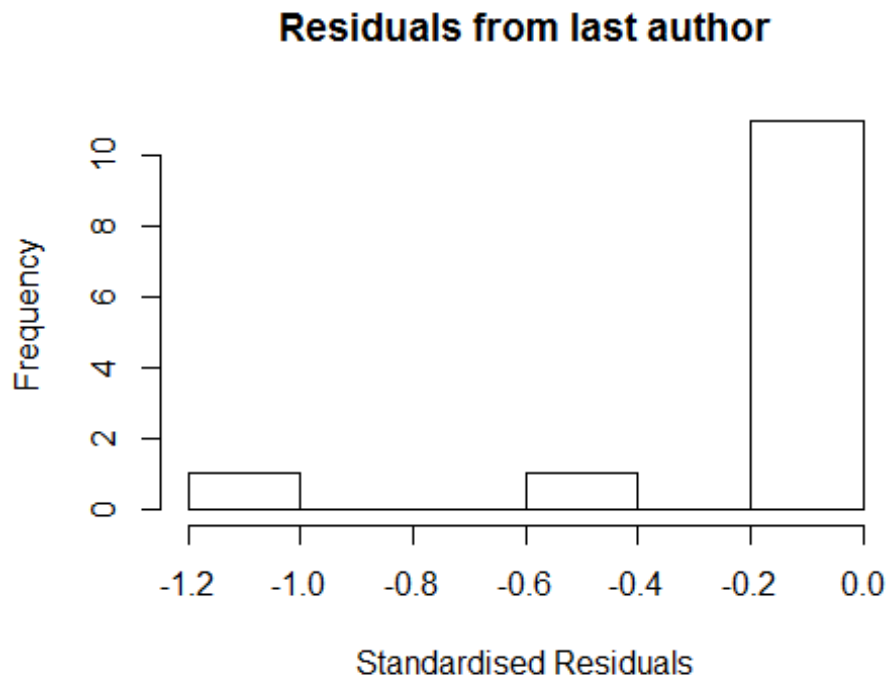
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 12, 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 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
## LastAuthorFemale  NaN 1          NaN
## UniqueAuthors    NaN 4          NaN
## Year              NaN 10         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 +
##       UniqueAuthors +
##       Year, data = AllScopusDataOlderFirstLastGendered, control =
##       lmrob.control(fast.s.large.n = Inf,
##       k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.87236 -0.25768  0.00397  0.23834  0.71651
##
```

```

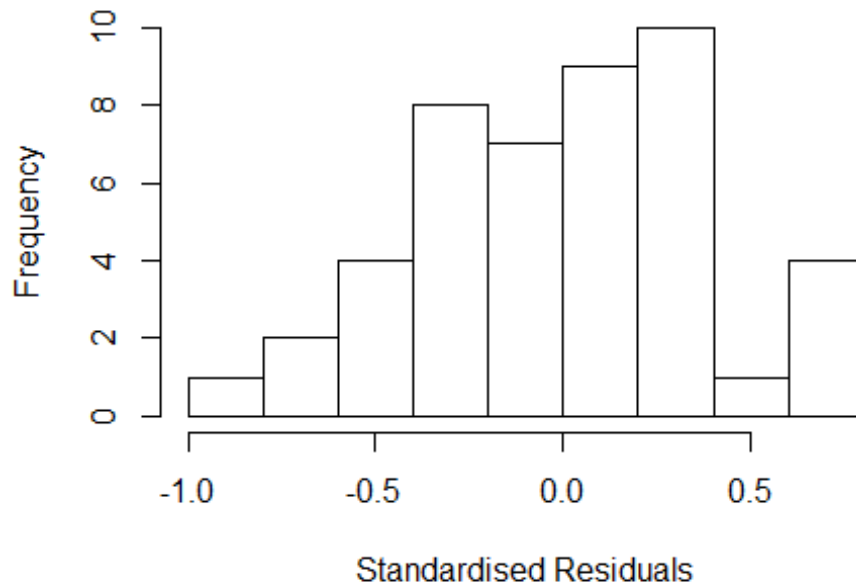
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.4143    0.2759   5.13 1.8e-05 ***
## FirstAuthorFemale1 -0.4818    0.1733  -2.78 0.00944 **
## LastAuthorFemale1  0.1719    0.1362   1.26 0.21705
## UniqueAuthors2     0.4775    0.2195   2.17 0.03795 *
## UniqueAuthors3     0.3960    0.2310   1.71 0.09717 .
## UniqueAuthors4     1.0220    0.2976   3.43 0.00181 **
## UniqueAuthors5     0.5499    0.2035   2.70 0.01140 *
## Year2003          -0.4950    0.0000  -Inf < 2e-16 ***
## Year2004          -0.1864    0.2449  -0.76 0.45261
## Year2005          -0.6076    0.1795  -3.38 0.00206 **
## Year2006          -1.4034    0.2563  -5.48 6.8e-06 ***
## Year2007          -1.2594    0.2125  -5.93 1.9e-06 ***
## Year2008          -1.7001    0.3817  -4.45 0.00012 ***
## Year2009          -0.0352    0.3292  -0.11 0.91560
## Year2010          -0.7489    0.2566  -2.92 0.00673 **
## Year2011          -0.5386    0.2380  -2.26 0.03127 *
## Year2012          -0.4798    0.2894  -1.66 0.10806
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.479
## Multiple R-squared:  0.625, Adjusted R-squared:  0.418
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~ = 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.720  0.920  0.961  0.935  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           2.17e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful
```

Residuals from first and last author and team size



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## LastAuthorFemale  NaN 1             NaN
## Year              NaN 10            NaN

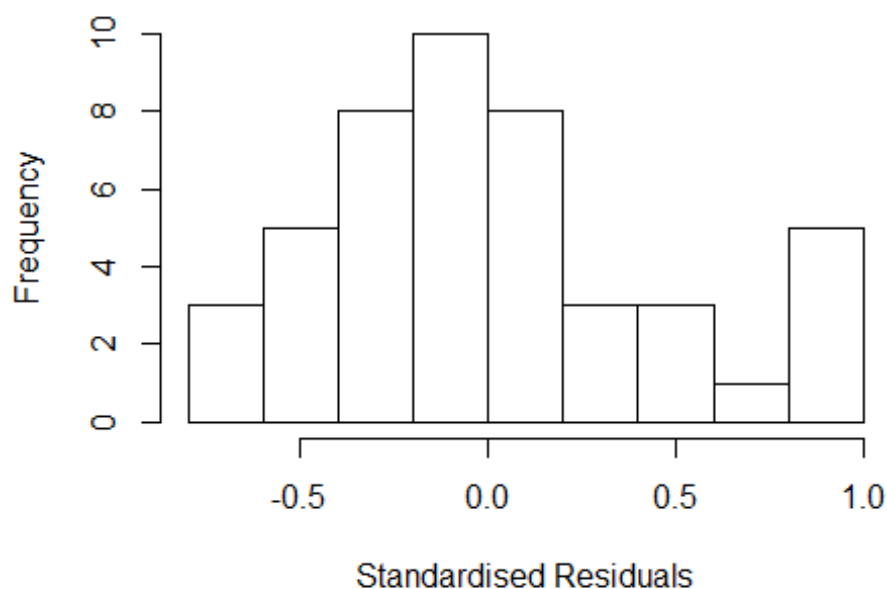
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId  NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
##       control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.7629 -0.3082 -0.0109  0.2694  0.9905
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      1.705      0.170   10.03 1.5e-11 ***
```

```

## FirstAuthorFemale1    -0.295      0.170    -1.74    0.09205 .
## LastAuthorFemale1     0.131      0.170     0.77    0.44451
## Year2003               -0.495      0.000    -Inf    < 2e-16 ***
## Year2004               -0.623      0.170    -3.67    0.00084 ***
## Year2005               -0.646      0.208    -3.11    0.00381 **
## Year2006               -1.224      0.189    -6.49    2.3e-07 ***
## Year2007               -1.172      0.175    -6.70    1.3e-07 ***
## Year2008               -1.333      0.332    -4.02    0.00032 ***
## Year2009                0.149      0.481     0.31    0.75882
## Year2010               -0.543      0.195    -2.78    0.00892 **
## Year2011               -0.352      0.309    -1.14    0.26296
## Year2012               -0.424      0.240    -1.77    0.08656 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.544
## Multiple R-squared:  0.508, Adjusted R-squared:  0.33
## Convergence in 12 IRWLS iterations
##
## Robustness weights:
## 7 weights are ~= 1. The remaining 39 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.721  0.919  0.959  0.934  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           2.17e-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"
## 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

```


Residuals from first and last author



```
##              GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale NaN 1             NaN
## Year              NaN 10            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
## -0.7752 -0.3016 -0.0101  0.2119  0.9460
##
## Coefficients:
##              Estimate Std. Error  t value Pr(>|t|)
## (Intercept)    1.69e+00   1.67e-01  1.01e+01  8.4e-12 ***
## FirstAuthorFemale1 -2.82e-01   1.67e-01 -1.69e+00  0.10033
## Year2003         -4.95e-01   0.00e+00    -Inf < 2e-16 ***
## Year2004         -4.92e-01   1.63e-08 -3.02e+07 < 2e-16 ***
## Year2005         -6.05e-01   2.41e-01 -2.51e+00  0.01681 *
## Year2006        -1.12e+00   9.89e-02 -1.13e+01  4.2e-13 ***
## Year2007        -1.10e+00   1.69e-01 -6.51e+00  1.9e-07 ***
## Year2008        -1.24e+00   3.21e-01 -3.86e+00  0.00048 ***
```

```

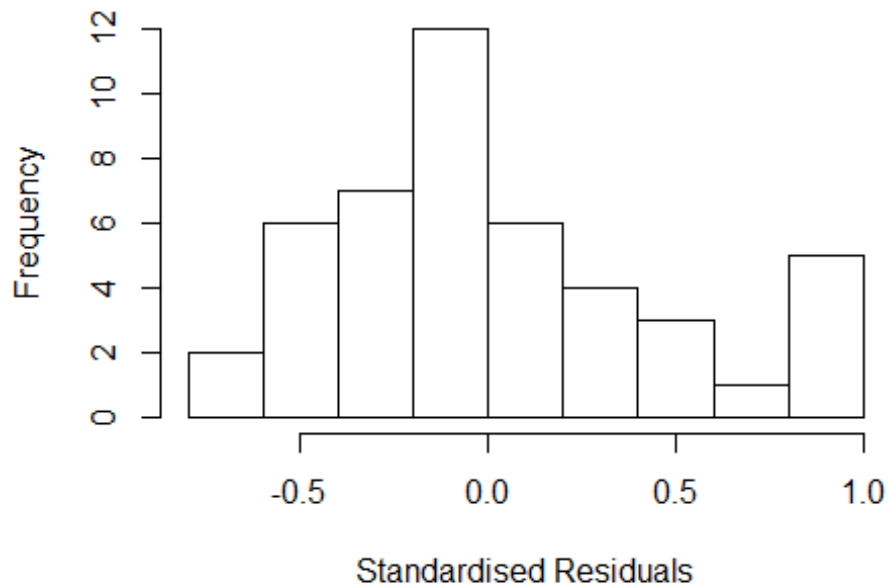
## Year2009          2.56e-01   3.58e-01   7.10e-01   0.47972
## Year2010          -4.58e-01   1.70e-01  -2.69e+00   0.01102 *
## Year2011          -2.53e-01   2.54e-01  -1.00e+00   0.32619
## Year2012          -3.30e-01   1.45e-01  -2.28e+00   0.02890 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.532
## Multiple R-squared:  0.506, Adjusted R-squared:  0.346
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 38 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.733  0.900  0.964  0.928  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      2.17e-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

## Warning in lf.cov(init, x = x): diag(.) had 0 or NA entries; non-finite
## result is doubtful

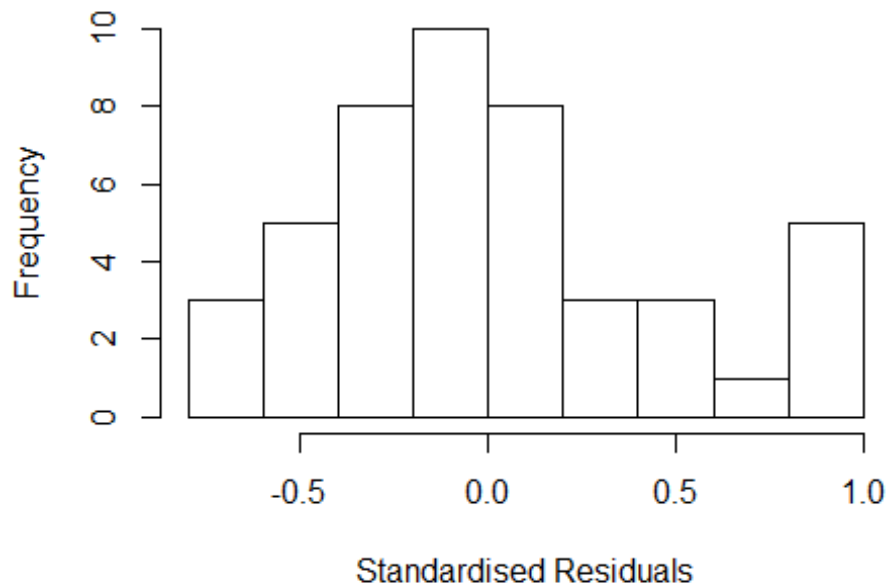
```

Residuals from first author



```
##          GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale NaN  1          NaN
## Year            NaN 10          NaN
```

Residuals from last author



```

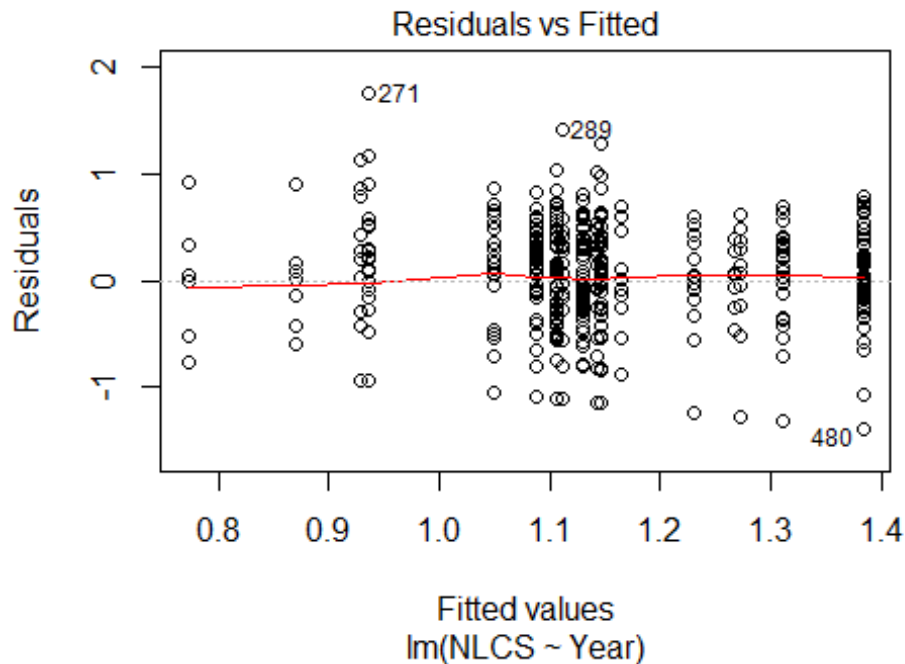
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -0.89971 -0.34049 -0.00204 0.27409 0.98829
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.410 0.000 Inf < 2e-16 ***
## LastAuthorFemale1 0.102 0.172 0.59 0.5576
## Year2003 -0.495 0.000 -Inf < 2e-16 ***
## Year2004 -0.594 0.172 -3.45 0.0015 **
## Year2005 -0.637 0.214 -2.98 0.0053 **
## Year2006 -1.134 0.229 -4.95 2.0e-05 ***
## Year2007 -1.068 0.162 -6.57 1.6e-07 ***
## Year2008 -1.143 0.253 -4.52 7.1e-05 ***
## Year2009 0.302 0.390 0.77 0.4447
## Year2010 -0.373 0.155 -2.41 0.0214 *
## Year2011 -0.186 0.285 -0.65 0.5181
## Year2012 -0.303 0.174 -1.74 0.0901 .
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.549
## Multiple R-squared: 0.474, Adjusted R-squared: 0.304
## Convergence in 11 IRWLS iterations
##
## Robustness weights:
## 13 weights are ~= 1. The remaining 33 ones are summarized as
## Min. 1st Qu. Median Mean 3rd Qu. Max.
## 0.726 0.867 0.951 0.916 0.975 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 2.17e-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: 46"
## [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
##   19   16   17   17   21   22   11   23   32   34   26   51   39   55   59
## 2011 2012
##   81   65
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   14   10    7   14    7   11    6   16   23   23   20   45   25   46   48
## 2011 2012
##   61   57
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010
##   14   10    7   14    7   11    5   14   22   21   20   43   23   39   41
## 2011 2012
##   53   51
## [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

```



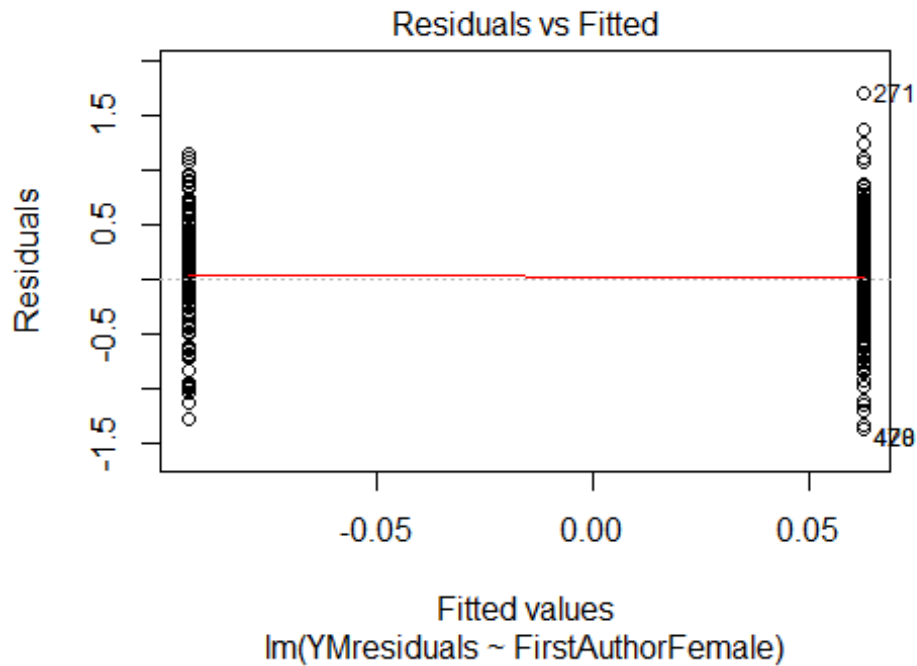
```
##
## Bartlett test of homogeneity of variances
##
## data: YMresiduals by FirstAuthorFemale
## Bartlett's K-squared = 0.79, df = 1, p-value = 0.4

## [1] "Female first author team size 2018 geometric mean: 3.4362434098782"
## [1] "Male first author team size 2018 geometric mean: 3.58427927026835"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

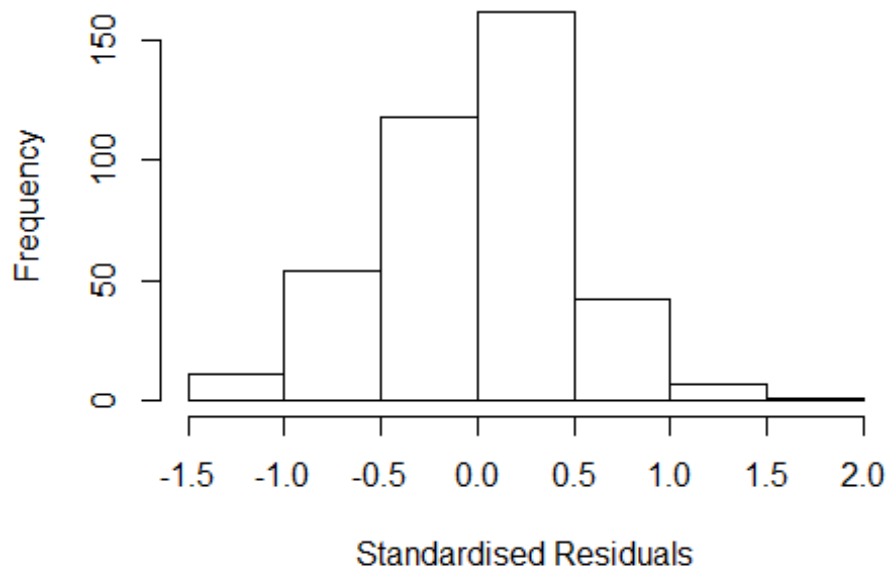
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 210, p-value = 0.9
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.24351597244694"
## [1] "Male last author team size 2018 geometric mean: 3.95988379152038"

## 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 = 74, 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 1.353 1          1.163
## LastAuthorFemale  1.339 1          1.157
## UniqueAuthors    2.677 4          1.131
## Year              3.424 16         1.039
```

Residuals from first and last author and team size



```
## [1] "List of 0 outliers with 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
## -1.4759 -0.3130 0.0392 0.3120 1.8157
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.8506 0.2357 3.61 0.00035 ***
## FirstAuthorFemale1 -0.1391 0.0579 -2.40 0.01680 *
## LastAuthorFemale1 -0.0629 0.0643 -0.98 0.32881
## UniqueAuthors2 0.2489 0.0812 3.07 0.00232 **
## UniqueAuthors3 0.2916 0.0800 3.65 0.00030 ***
## UniqueAuthors4 0.2530 0.0950 2.66 0.00810 **
## UniqueAuthors5 0.4373 0.0848 5.15 4.1e-07 ***
## Year1997 0.1202 0.2587 0.46 0.64257
## Year1998 0.3495 0.2683 1.30 0.19347
## Year1999 0.2396 0.2460 0.97 0.33065
```

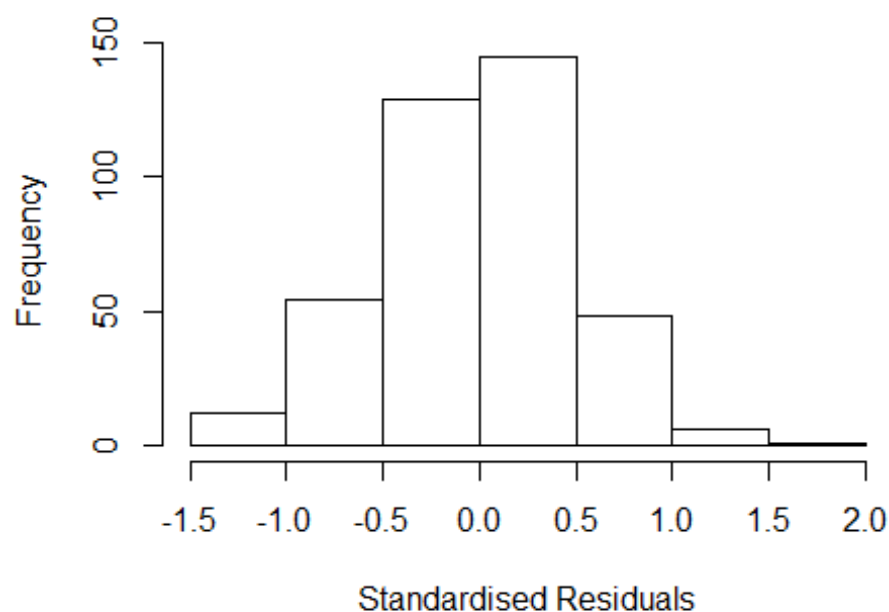


```

## Year2000          -0.1193      0.2915    -0.41  0.68263
## Year2001           0.3113      0.2506     1.24  0.21499
## Year2002          -0.3148      0.4213    -0.75  0.45539
## Year2003           0.2036      0.2700     0.75  0.45123
## Year2004           0.1156      0.2665     0.43  0.66471
## Year2005          -0.1623      0.2935    -0.55  0.58044
## Year2006           0.0807      0.2622     0.31  0.75844
## Year2007           0.1270      0.2419     0.52  0.59991
## Year2008           0.3966      0.2521     1.57  0.11654
## Year2009           0.3396      0.2430     1.40  0.16305
## Year2010           0.1418      0.2449     0.58  0.56296
## Year2011           0.1529      0.2369     0.65  0.51913
## Year2012           0.0569      0.2407     0.24  0.81319
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.469
## Multiple R-squared:  0.166, Adjusted R-squared:  0.116
## Convergence in 16 IRWLS iterations
##
## Robustness weights:
## 31 weights are ~= 1. The remaining 364 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.101  0.864  0.952   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      2.53e-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 1.335 1      1.155
## LastAuthorFemale  1.271 1      1.127
## Year              1.360 16      1.010

```

Residuals from first and last author



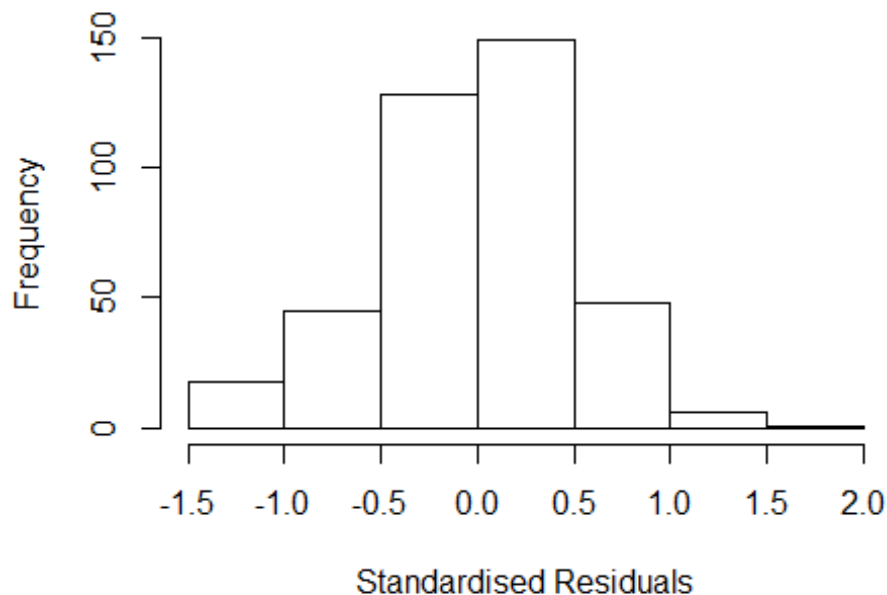
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.49272 -0.32622  0.00759  0.31648  1.88545
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9818    0.2307   4.26 2.6e-05 ***
## FirstAuthorFemale1 -0.1231    0.0599  -2.06  0.041 *
## LastAuthorFemale1  -0.1072    0.0661  -1.62  0.105
## Year1997           0.2029    0.2765   0.73  0.463
## Year1998           0.3572    0.2491   1.43  0.152
## Year1999           0.3305    0.2547   1.30  0.195
## Year2000          -0.0351    0.2711  -0.13  0.897
## Year2001           0.3937    0.2599   1.52  0.131
## Year2002          -0.2194    0.3678  -0.60  0.551
## Year2003           0.2323    0.2833   0.82  0.413
## Year2004           0.2405    0.2665   0.90  0.367
## Year2005          -0.0700    0.3001  -0.23  0.816
```

```

## Year2006          0.2223      0.2566      0.87      0.387
## Year2007          0.2470      0.2420      1.02      0.308
## Year2008          0.5110      0.2536      2.01      0.045 *
## Year2009          0.4584      0.2414      1.90      0.058 .
## Year2010          0.2536      0.2451      1.03      0.301
## Year2011          0.2441      0.2395      1.02      0.309
## Year2012          0.2066      0.2401      0.86      0.390
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.486
## Multiple R-squared:  0.106, Adjusted R-squared:  0.0637
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 37 weights are ~= 1. The remaining 358 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0983 0.8610 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      2.53e-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 1.198 1      1.095
## Year              1.198 16      1.006

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.4821 -0.3367 0.0172 0.3122 1.7710
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 0.9739 0.2317 4.20 3.3e-05 ***
## FirstAuthorFemale1 -0.1632 0.0576 -2.83 0.0048 **
## Year1997 0.2150 0.2762 0.78 0.4369
## Year1998 0.3448 0.2499 1.38 0.1686
## Year1999 0.3407 0.2563 1.33 0.1846
## Year2000 -0.0555 0.2795 -0.20 0.8427
## Year2001 0.4006 0.2626 1.53 0.1280
## Year2002 -0.2729 0.3874 -0.70 0.4817
## Year2003 0.2310 0.2859 0.81 0.4196
## Year2004 0.2085 0.2641 0.79 0.4303
## Year2005 -0.0549 0.3068 -0.18 0.8582
## Year2006 0.1992 0.2560 0.78 0.4369
```

```

## Year2007          0.2341      0.2428      0.96      0.3356
## Year2008          0.5082      0.2528      2.01      0.0451 *
## Year2009          0.4499      0.2420      1.86      0.0638 .
## Year2010          0.2421      0.2456      0.99      0.3249
## Year2011          0.2394      0.2401      1.00      0.3194
## Year2012          0.1919      0.2406      0.80      0.4255
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.484
## Multiple R-squared:  0.0989, Adjusted R-squared:  0.0582
## Convergence in 17 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 357 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.151  0.867  0.952  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      2.53e-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 1.145 1      1.070
## Year      1.145 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
## -1.44673 -0.32240  0.00181  0.33705  1.98426

```

```

##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)      0.9541    0.2244   4.25 2.7e-05 ***
## LastAuthorFemale1 -0.1606    0.0645  -2.49  0.013 *
## Year1997          0.2189    0.2754   0.79  0.427
## Year1998          0.3636    0.2445   1.49  0.138
## Year1999          0.3311    0.2497   1.33  0.186
## Year2000         -0.0414    0.2671  -0.15  0.877
## Year2001          0.4000    0.2556   1.56  0.118
## Year2002         -0.1884    0.3453  -0.55  0.586
## Year2003          0.2359    0.2808   0.84  0.401
## Year2004          0.2566    0.2622   0.98  0.328
## Year2005         -0.0878    0.2939  -0.30  0.765
## Year2006          0.2304    0.2559   0.90  0.368
## Year2007          0.2380    0.2363   1.01  0.315
## Year2008          0.4926    0.2469   2.00  0.047 *
## Year2009          0.4576    0.2351   1.95  0.052 .
## Year2010          0.2565    0.2402   1.07  0.286
## Year2011          0.2239    0.2337   0.96  0.339
## Year2012          0.2111    0.2356   0.90  0.371
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.487
## Multiple R-squared:  0.0967, Adjusted R-squared:  0.056
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 38 weights are ~= 1. The remaining 357 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
## 0.0595 0.8660 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      2.53e-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: 395"
## [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]"
##
## 2011
## 1
##
## 2011
## 1
##
## 2011
## 1
## [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"
## [1] "Regression 4: Last author gender, Year as factors"
## [1] "Sample size for the above analysis: 1"
## [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 2010 2011
## 2 3 3 3 5 2 2 3 2 2 5 6 3 2 4
## 2012
## 4
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011
## 0 1 2 3 2 0 2 3 1 1 3 5 1 1 2
## 2012
## 2
##
## 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2010 2011
## 0 0 2 3 2 0 2 3 0 1 3 5 1 1 2
## 2012
## 2
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"

```

```
## [1] "Female first author team size 2018 geometric mean: 2"
## [1] "Male first author team size 2018 geometric mean: 2.88449914061482"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

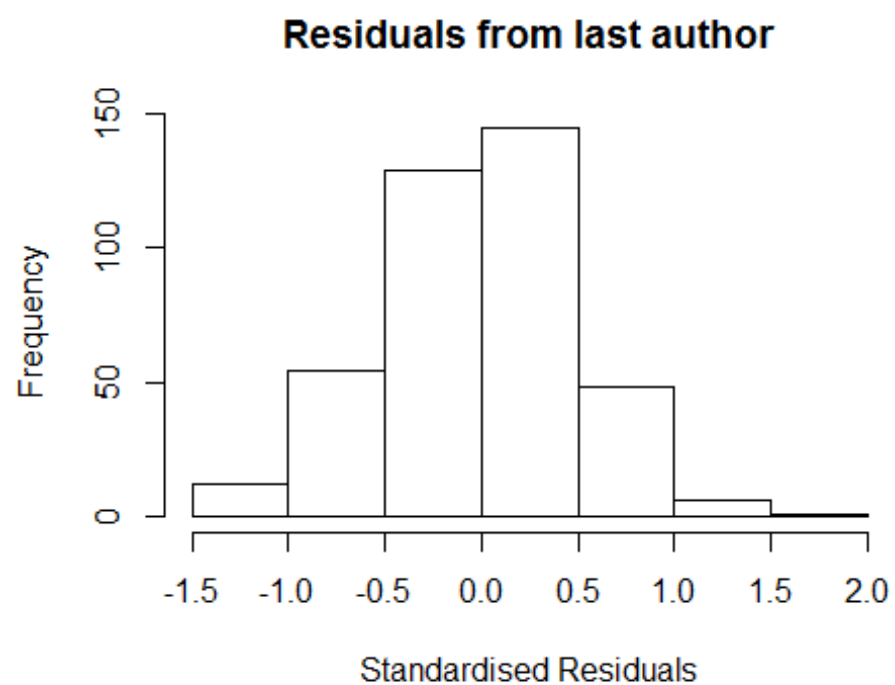
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2"
## [1] "Male last author team size 2018 geometric mean: 2.88449914061482"

## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties

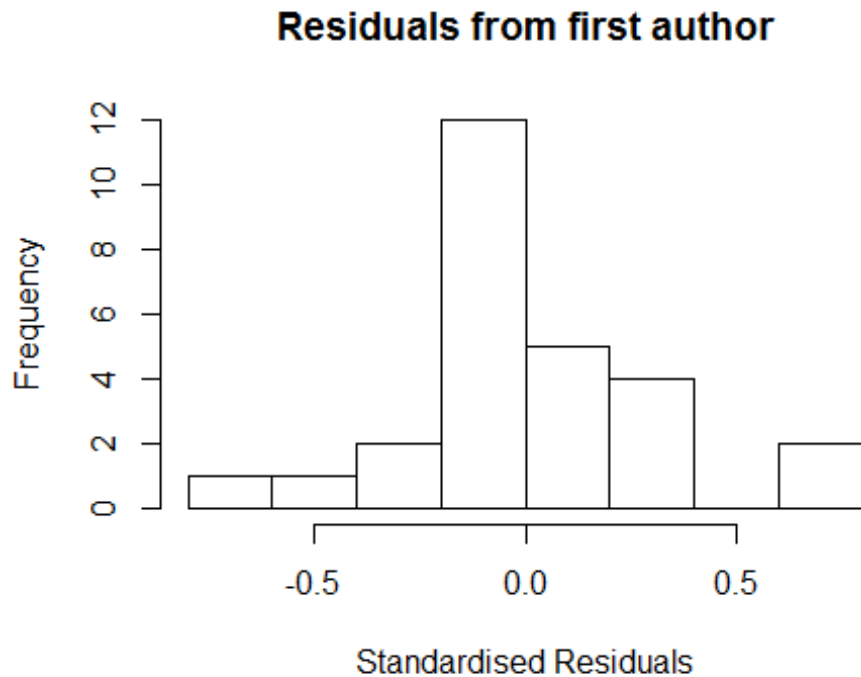
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 0.5, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Regression 1: First author gender, last author gender, team size,
Year as factors"
## [1] "Regression 2: First author gender, Last author gender, Year as
factors"

## Warning in lmrob.S(x, y, control = control, mf = mf): 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 3: First author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## FirstAuthorFemale 43.57  1          6.600
## Year              43.57 11          1.187
```



```
## [1] "Regression 4: Last author gender, Year as factors"
##           GVIF Df GVIF^(1/(2*Df))
## LastAuthorFemale 6.56 1      2.561
## Year             6.56 11      1.089

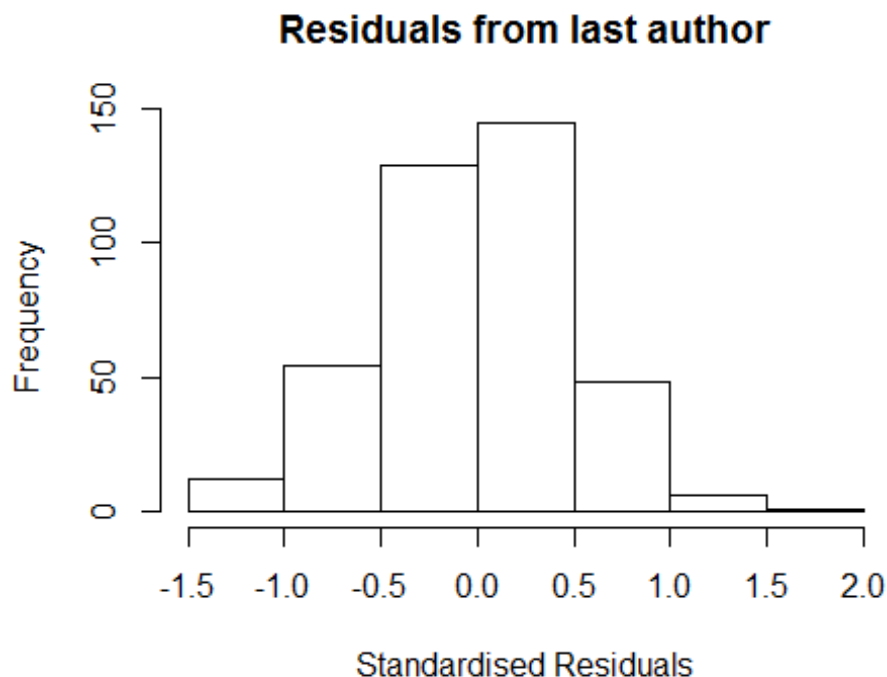
## [1] "Sample size for the above analysis: 27"
## [1] ""
## [1] ""
## [1] "#####"
## [1] "Analysis of AJSC 3616"
## [1] "#####"
## [1] "Sample sizes for all years [All, first gendered, first & last
gendered] [check that these decrease]"
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    1    1    1    4    4    3    1    5    7   12   10    7   11    9
## 2012
##   16
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    1    1    1    3    3    2    1    5    7   11    9    6   10    9
## 2012
##   13
##
## 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011
##    2    0    1    1    2    3    2    1    5    6   11    8    6    8    9
## 2012
```

```
## 12
## [1] "Heteroscedasticity checks, confirming that there are problems with
these"
## [1] "Female first author team size 2018 geometric mean: 2.95156811701606"
## [1] "Male first author team size 2018 geometric mean: 2.40224886796286"

## 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 = 24, p-value = 0.6
## alternative hypothesis: true location shift is not equal to 0
##
## [1] "Female last author team size 2018 geometric mean: 2.94471340036069"
## [1] "Male last author team size 2018 geometric mean: 2.49288287167843"

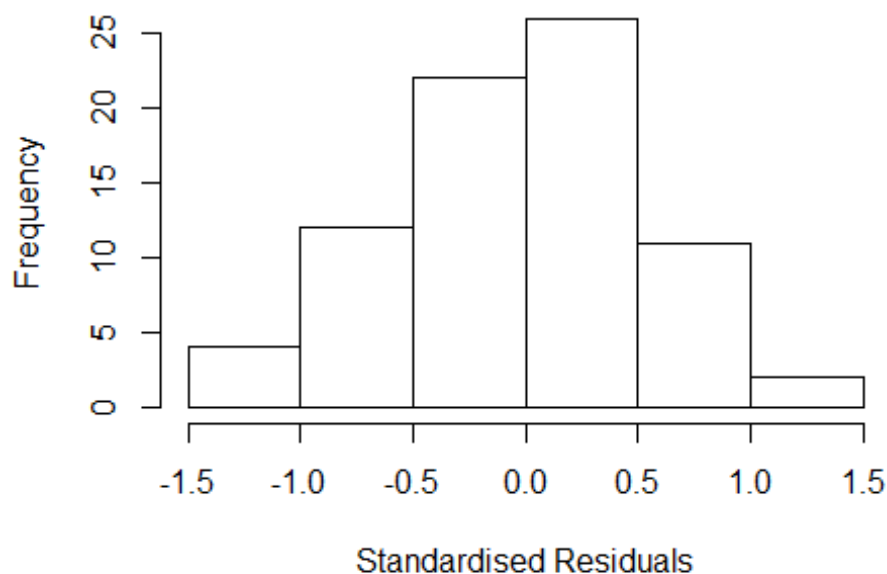
## Warning in wilcox.test.default(FemaleTeamSizes2018, MaleTeamSizes2018,
## alternative = "two.sided"): cannot compute exact p-value with ties
```



```
##
## Wilcoxon rank sum test with continuity correction
##
## data: FemaleTeamSizes2018 and MaleTeamSizes2018
## W = 24, 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 1.393e+14  1      1.180e+07
## LastAuthorFemale  2.629e+00  1      1.621e+00
## UniqueAuthors    6.450e+28  4      3.992e+03
## Year              5.293e+29 14      1.152e+01
```

Residuals from first and last author and team size



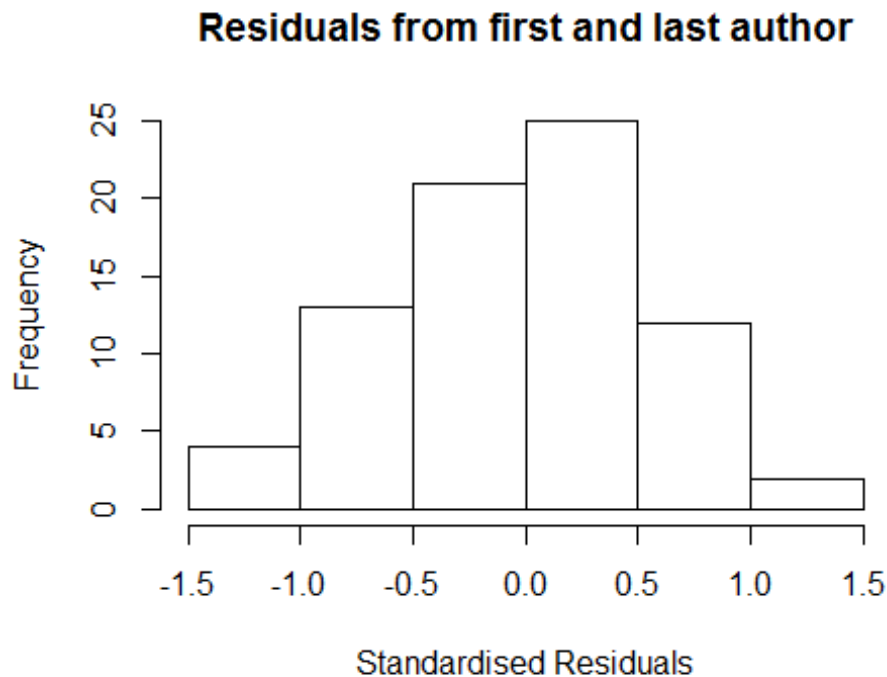
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS      Year      OneField  Fields    residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale +
UniqueAuthors +
##      Year, data = AllScopusDataOlderFirstLastGendered, control =
lmrob.control(fast.s.large.n = Inf,
##      k.max = 1000))
## \--> method = "MM"
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.2263 -0.3927  0.0233  0.3838  1.4650
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```

## (Intercept)          1.2109      0.8083      1.50      0.14
## FirstAuthorFemale1 -0.1841      0.2693     -0.68      0.50
## LastAuthorFemale1   0.2014      0.2199      0.92      0.36
## UniqueAuthors2     -0.0132      0.2694     -0.05      0.96
## UniqueAuthors3     -0.0334      0.3012     -0.11      0.91
## UniqueAuthors4       0.4254      0.3211      1.33      0.19
## UniqueAuthors5     -0.2244      1.1023     -0.20      0.84
## Year1999            -0.7729      0.8345     -0.93      0.36
## Year2000             0.9408      0.8355      1.13      0.26
## Year2001            -0.5062      1.0976     -0.46      0.65
## Year2002             0.4575      0.8650      0.53      0.60
## Year2003            -0.1088      0.8730     -0.12      0.90
## Year2004            -1.0271      0.8083     -1.27      0.21
## Year2005             0.2329      0.8684      0.27      0.79
## Year2006             0.2195      0.8527      0.26      0.80
## Year2007             0.0261      0.8141      0.03      0.97
## Year2008             0.1819      0.8379      0.22      0.83
## Year2009            -0.2958      0.8395     -0.35      0.73
## Year2010             0.1207      0.8668      0.14      0.89
## Year2011            -0.4509      0.9359     -0.48      0.63
## Year2012            -0.1187      0.8675     -0.14      0.89
##
## Robust residual standard error: 0.6
## Multiple R-squared:  0.221, Adjusted R-squared:  -0.0567
## Convergence in 25 IRWLS iterations
##
## Robustness weights:
## 8 weights are ~= 1. The remaining 69 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.530  0.863   0.951   0.915  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      1.30e-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 1.338e+14 1      1.157e+07

```

## LastAuthorFemale	1.910e+00	1	1.382e+00
## Year	2.348e+14	14	3.260e+00



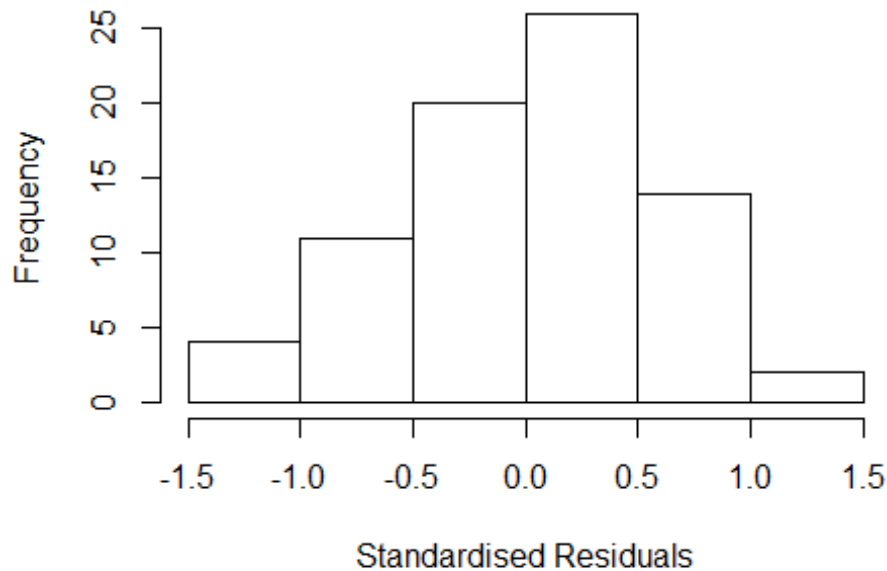
```
## [1] "List of 0 outliers with residuals above 2.5"
## [1] ScopusId NLCS Year OneField Fields residuals
## <0 rows> (or 0-length row.names)
##
## Call:
## lmrob(formula = NLCS ~ FirstAuthorFemale + LastAuthorFemale + Year, data =
AllScopusDataOlderFirstLastGendered,
## control = lmrob.control(fast.s.large.n = Inf, k.max = 1000))
## \--> method = "MM"
## Residuals:
## Min 1Q Median 3Q Max
## -1.25106 -0.40530 0.00176 0.41347 1.13043
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2214 0.7031 1.74 0.088 .
## FirstAuthorFemale1 -0.1963 0.1878 -1.05 0.300
## LastAuthorFemale1 0.1672 0.2042 0.82 0.416
## Year1999 -0.7826 0.7055 -1.11 0.272
## Year2000 0.9767 0.7145 1.37 0.177
## Year2001 -0.2747 1.2425 -0.22 0.826
## Year2002 0.3979 0.7390 0.54 0.592
## Year2003 -0.1128 0.7687 -0.15 0.884
## Year2004 -1.0166 0.7055 -1.44 0.155
```

```

## Year2005          0.2190      0.7759      0.28      0.779
## Year2006          0.2243      0.7394      0.30      0.763
## Year2007          0.0889      0.7169      0.12      0.902
## Year2008          0.2423      0.7148      0.34      0.736
## Year2009         -0.2100      0.7317     -0.29      0.775
## Year2010          0.1831      0.7559      0.24      0.809
## Year2011         -0.3171      0.7701     -0.41      0.682
## Year2012          0.0587      0.7392      0.08      0.937
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.68
## Multiple R-squared:  0.157, Adjusted R-squared:  -0.0673
## Convergence in 15 IRWLS iterations
##
## Robustness weights:
## 10 weights are ~= 1. The remaining 67 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.715  0.877   0.956   0.926   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      1.30e-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 -4.014e+28 1      NaN
## Year              -4.014e+28 14      NaN

```

Residuals from first author



```
## [1] "List of 0 outliers 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
## -1.4019 -0.3658 0.0325 0.3912 1.1885
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.3050 0.5874 2.22 0.03 *
## FirstAuthorFemale1 -0.1245 0.1700 -0.73 0.47
## Year1999 -0.6990 0.5874 -1.19 0.24
## Year2000 0.9885 0.6090 1.62 0.11
## Year2001 -0.3465 1.0482 -0.33 0.74
## Year2002 0.3748 0.6400 0.59 0.56
## Year2003 -0.1488 0.6526 -0.23 0.82
## Year2004 -0.9330 0.5874 -1.59 0.12
## Year2005 0.2213 0.7073 0.31 0.76
## Year2006 0.2048 0.6419 0.32 0.75
## Year2007 0.0789 0.6141 0.13 0.90
## Year2008 0.2218 0.6074 0.37 0.72
```

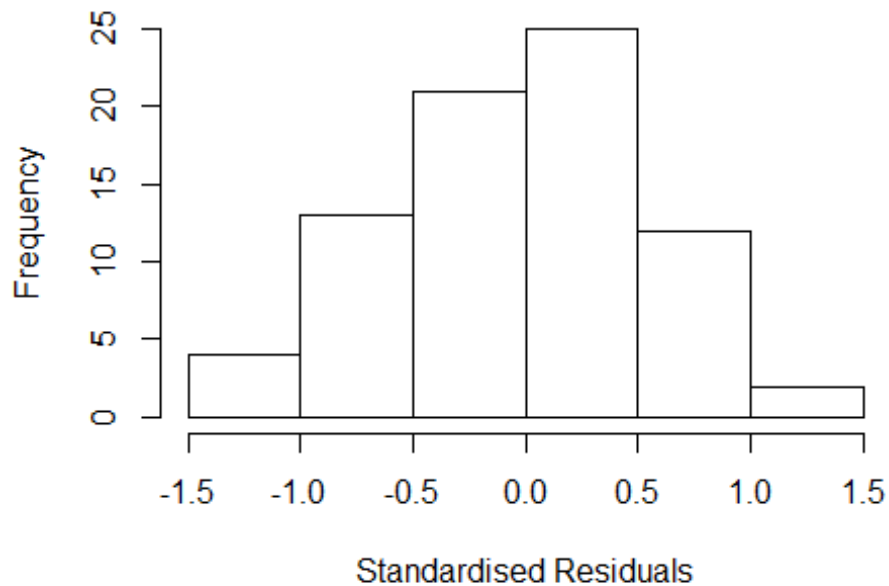


```

## Year2009          -0.2037      0.6327   -0.32      0.75
## Year2010           0.1817      0.6594     0.28      0.78
## Year2011          -0.2915      0.6760   -0.43      0.67
## Year2012           0.0255      0.6345     0.04      0.97
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.663
## Multiple R-squared:  0.152, Adjusted R-squared:  -0.0571
## Convergence in 13 IRWLS iterations
##
## Robustness weights:
## 11 weights are ~= 1. The remaining 66 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.634  0.873   0.952   0.920   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      1.30e-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 1.511  1      1.229
## Year            1.511 14      1.015

```

Residuals from last author



```
## [1] "List of 0 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
## -1.4053 -0.4297 0.0317 0.4254 1.2884
##
## Coefficients:
## Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.2622 0.6551 1.93 0.059 .
## LastAuthorFemale1 0.0857 0.1854 0.46 0.646
## Year1999 -0.7418 0.6527 -1.14 0.260
## Year2000 0.8212 0.6527 1.26 0.213
## Year2001 -0.4710 1.1497 -0.41 0.683
## Year2002 0.2839 0.6863 0.41 0.681
## Year2003 -0.2110 0.7444 -0.28 0.778
## Year2004 -0.9758 0.6527 -1.50 0.140
## Year2005 0.1432 0.7499 0.19 0.849
## Year2006 0.1062 0.6819 0.16 0.877
## Year2007 -0.0309 0.6629 -0.05 0.963
## Year2008 0.0911 0.6543 0.14 0.890
```

```

## Year2009          -0.2578      0.6783    -0.38      0.705
## Year2010           0.0449      0.7064      0.06      0.950
## Year2011          -0.4343      0.7305    -0.59      0.554
## Year2012          -0.0622      0.6807    -0.09      0.927
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Robust residual standard error: 0.662
## Multiple R-squared:  0.149, Adjusted R-squared:  -0.0601
## Convergence in 14 IRWLS iterations
##
## Robustness weights:
## 9 weights are ~= 1. The remaining 68 ones are summarized as
##   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  0.632  0.888  0.956  0.922  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      1.30e-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: 77"

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"

```

```
print(BasicStats)
```

##	1Ctry	Gend	1st	nth	Fem1	96	Fem1	18	Chg1	FemN	96	FemN	18	ChgN
##	1000	583		49.40	16.667	34.884	18.2171	16.667	27.907	11.2403				
##	1100	1399		58.11	15.385	57.377	41.9924	11.538	27.869	16.3304				
##	1101	254		48.82	16.667	37.500	20.8333	0.000	37.500	37.5000				
##	1102	3160		23.73	26.923	36.364	9.4406	11.538	40.260	28.7213				
##	1103	3640		32.42	15.625	43.678	28.0532	12.500	36.782	24.2816				
##	1104	2929		53.81	20.000	42.029	22.0290	18.182	21.739	3.5573				
##	1105	5578		55.61	25.000	40.580	15.5797	19.737	30.435	10.6979				
##	1106	2545		37.80	35.484	48.387	12.9032	32.258	33.871	1.6129				
##	1107	902		48.67	0.000	13.333	13.3333	7.692	26.667	18.9744				
##	1108	1484		21.43	23.810	48.387	24.5776	14.286	45.161	30.8756				
##	1109	1282		34.95	20.000	50.000	30.0000	10.000	40.909	30.9091				
##	1110	3339		35.37	17.021	42.188	25.1662	21.277	32.812	11.5359				
##	1111	1585		20.13	12.500	34.483	21.9828	0.000	31.034	31.0345				
##	1200	103		79.61	NaN	25.000	NaN	NaN	75.000	NaN				
##	1201	1005		74.53	60.000	58.140	-1.8605	60.000	58.140	-1.8605				
##	1202	912		81.14	21.429	45.000	23.5714	21.429	38.333	16.9048				
##	1203	968		76.45	81.818	54.054	-27.7641	72.727	62.162	-10.5651				
##	1204	205		66.83	33.333	22.222	-11.1111	33.333	11.111	-22.2222				
##	1205	53		84.91	NaN	100.000	NaN	NaN	100.000	NaN				
##	1206	35		74.29	NaN	100.000	NaN	NaN	80.000	NaN				
##	1207	376		83.24	57.143	45.833	-11.3095	71.429	50.000	-21.4286				
##	1208	542		85.98	0.000	55.556	55.5556	0.000	44.444	44.4444				
##	1209	24		70.83	NaN	100.000	NaN	NaN	100.000	NaN				
##	1210	110		79.09	0.000	60.000	60.0000	0.000	40.000	40.0000				
##	1211	736		83.29	19.048	33.333	14.2857	19.048	25.641	6.5934				
##	1212	247		88.26	NaN	23.529	NaN	NaN	17.647	NaN				
##	1213	436		82.57	NaN	64.516	NaN	NaN	54.839	NaN				
##	1300	1194		63.57	33.333	57.143	23.8095	0.000	30.357	30.3571				
##	1301	12		66.67	NaN	0.000	NaN	NaN	100.000	NaN				
##	1302	81		58.02	0.000	83.333	83.3333	50.000	33.333	-16.6667				
##	1303	2399		59.69	27.778	46.939	19.1610	12.500	42.857	30.3571				
##	1304	610		53.61	33.333	55.556	22.2222	13.333	5.556	-7.7778				
##	1305	1139		51.27	35.294	44.737	9.4427	5.882	26.316	20.4334				
##	1306	362		55.80	33.333	37.500	4.1667	33.333	50.000	16.6667				
##	1307	954		55.56	34.783	56.522	21.7391	34.783	43.478	8.6957				
##	1308	680		61.47	28.571	66.667	38.0952	28.571	27.778	-0.7937				
##	1309	386		50.78	50.000	71.429	21.4286	66.667	42.857	-23.8095				
##	1310	850		44.35	13.333	75.000	61.6667	13.333	25.000	11.6667				
##	1311	1892		49.74	32.558	40.000	7.4419	16.279	27.500	11.2209				
##	1312	1796		60.02	30.612	45.455	14.8423	18.367	38.636	20.2690				
##	1313	424		66.27	0.000	46.667	46.6667	0.000	33.333	33.3333				
##	1314	1213		53.59	17.857	50.000	32.1429	10.714	32.353	21.6387				
##	1315	239		59.00	25.000	40.000	15.0000	12.500	40.000	27.5000				
##	1400	539		70.69	45.455	27.273	-18.1818	45.455	18.182	-27.2727				
##	1401	265		70.19	28.571	40.000	11.4286	28.571	40.000	11.4286				
##	1402	543		78.27	25.000	33.333	8.3333	8.333	37.037	28.7037				
##	1403	858		70.86	44.444	38.710	-5.7348	44.444	35.484	-8.9606				

## 1404	167	49.70	0.000	30.000	30.0000	0.000	60.000	60.0000
## 1405	493	77.89	21.569	22.222	0.6536	21.569	27.778	6.2092
## 1406	610	71.97	0.000	53.571	53.5714	0.000	50.000	50.0000
## 1407	536	77.24	41.667	33.333	-8.3333	41.667	33.333	-8.3333
## 1408	1037	71.36	28.000	43.182	15.1818	36.000	50.000	14.0000
## 1409	761	77.79	0.000	47.826	47.8261	18.182	56.522	38.3399
## 1410	80	80.00	100.000	25.000	-75.0000	100.000	0.000	-100.0000
## 1500	989	40.44	0.000	21.053	21.0526	12.500	10.526	-1.9737
## 1501	57	47.37	NaN	0.000	NaN	NaN	100.000	NaN
## 1502	351	39.60	33.333	33.333	0.0000	0.000	11.111	11.1111
## 1503	272	64.71	14.286	16.667	2.3810	14.286	16.667	2.3810
## 1504	17	70.59	NaN	NaN	NaN	NaN	NaN	NaN
## 1505	111	48.65	NaN	0.000	NaN	NaN	50.000	NaN
## 1506	30	43.33	NaN	0.000	NaN	NaN	0.000	NaN
## 1507	88	30.68	NaN	20.000	NaN	NaN	0.000	NaN
## 1508	96	38.54	50.000	20.000	-30.0000	0.000	0.000	0.0000
## 1600	1655	55.41	12.821	27.778	14.9573	10.256	16.667	6.4103
## 1601	68	39.71	NaN	80.000	NaN	NaN	40.000	NaN
## 1602	646	52.79	33.333	25.000	-8.3333	16.667	18.750	2.0833
## 1603	164	37.80	NaN	100.000	NaN	NaN	0.000	NaN
## 1604	617	69.21	22.727	33.333	10.6061	18.182	44.444	26.2626
## 1605	1279	64.27	30.000	40.000	10.0000	13.333	35.000	21.6667
## 1606	1217	61.79	21.053	29.167	8.1140	13.158	29.167	16.0088
## 1607	277	54.87	0.000	25.000	25.0000	0.000	0.000	0.0000
## 1700	320	59.06	20.000	11.111	-8.8889	10.000	11.111	1.1111
## 1701	31	64.52	NaN	0.000	NaN	NaN	0.000	NaN
## 1702	325	45.54	20.000	25.000	5.0000	20.000	50.000	30.0000
## 1703	224	54.02	0.000	0.000	0.0000	0.000	50.000	50.0000
## 1704	83	48.19	0.000	0.000	0.0000	0.000	33.333	33.3333
## 1705	325	51.69	0.000	20.000	20.0000	0.000	0.000	0.0000
## 1706	995	47.14	60.000	13.333	-46.6667	40.000	13.333	-26.6667
## 1707	155	41.94	NaN	0.000	NaN	NaN	0.000	NaN
## 1708	162	43.21	NaN	100.000	NaN	NaN	0.000	NaN
## 1709	175	51.43	0.000	40.000	40.0000	0.000	20.000	20.0000
## 1710	379	53.30	66.667	30.769	-35.8974	33.333	15.385	-17.9487
## 1711	156	34.62	50.000	20.000	-30.0000	0.000	0.000	0.0000
## 1712	639	41.63	10.000	9.091	-0.9091	10.000	0.000	-10.0000
## 1800	110	63.64	NaN	33.333	NaN	NaN	33.333	NaN
## 1801	3	66.67	NaN	NaN	NaN	NaN	NaN	NaN
## 1802	195	64.10	66.667	28.571	-38.0952	33.333	42.857	9.5238
## 1803	332	55.42	33.333	11.111	-22.2222	33.333	22.222	-11.1111
## 1804	255	56.08	0.000	0.000	0.0000	0.000	0.000	0.0000
## 1900	622	61.58	21.429	35.000	13.5714	21.429	40.000	18.5714
## 1901	966	46.89	12.500	44.444	31.9444	6.250	14.815	8.5648
## 1902	533	52.72	0.000	16.667	16.6667	0.000	8.333	8.3333
## 1903	77	59.74	0.000	0.000	0.0000	0.000	0.000	0.0000
## 1904	725	55.45	21.429	23.333	1.9048	14.286	23.333	9.0476
## 1905	78	48.72	NaN	0.000	NaN	NaN	0.000	NaN
## 1906	769	55.53	0.000	45.000	45.0000	0.000	30.000	30.0000
## 1907	1126	51.07	19.048	50.000	30.9524	4.762	9.524	4.7619

## 1908	1223	49.47	8.333	40.000	31.6667	4.167	20.000	15.8333
## 1909	613	53.34	0.000	32.000	32.0000	0.000	4.000	4.0000
## 1910	842	58.55	12.500	37.500	25.0000	12.500	37.500	25.0000
## 1911	294	56.46	10.000	22.222	12.2222	10.000	33.333	23.3333
## 1912	377	45.89	0.000	80.000	80.0000	0.000	20.000	20.0000
## 1913	88	44.32	NaN	66.667	NaN	NaN	0.000	NaN
## 2000	395	76.71	14.286	5.000	-9.2857	14.286	30.000	15.7143
## 2001	172	73.84	0.000	10.000	10.0000	0.000	40.000	40.0000
## 2002	1367	68.91	15.000	22.414	7.4138	15.000	24.138	9.1379
## 2003	704	70.03	18.182	26.316	8.1340	18.182	28.947	10.7656
## 2100	179	51.40	NaN	50.000	NaN	NaN	16.667	NaN
## 2101	11	45.45	NaN	0.000	NaN	NaN	0.000	NaN
## 2102	259	36.29	0.000	25.000	25.0000	0.000	12.500	12.5000
## 2103	122	41.80	0.000	40.000	40.0000	0.000	20.000	20.0000
## 2104	54	40.74	0.000	50.000	50.0000	0.000	25.000	25.0000
## 2105	482	50.21	0.000	32.143	32.1429	0.000	28.571	28.5714
## 2200	397	46.85	23.077	11.111	-11.9658	7.692	11.111	3.4188
## 2201	102	69.61	NaN	33.333	NaN	NaN	0.000	NaN
## 2202	107	49.53	NaN	0.000	NaN	NaN	0.000	NaN
## 2203	80	47.50	NaN	50.000	NaN	NaN	50.000	NaN
## 2204	424	43.16	NaN	27.778	NaN	NaN	22.222	NaN
## 2205	969	47.47	0.000	28.571	28.5714	0.000	7.143	7.1429
## 2206	37	37.84	0.000	NaN	NaN	0.000	NaN	NaN
## 2207	330	30.61	0.000	0.000	0.0000	0.000	0.000	0.0000
## 2208	1224	38.40	0.000	12.500	12.5000	0.000	25.000	25.0000
## 2209	626	34.66	NaN	18.750	NaN	NaN	18.750	NaN
## 2210	983	30.21	0.000	17.391	17.3913	0.000	8.696	8.6957
## 2211	660	27.88	NaN	22.727	NaN	NaN	9.091	NaN
## 2212	130	57.69	0.000	20.000	20.0000	0.000	0.000	0.0000
## 2213	359	56.82	50.000	27.778	-22.2222	25.000	22.222	-2.7778
## 2214	119	62.18	0.000	0.000	0.0000	0.000	0.000	0.0000
## 2215	425	44.24	NaN	28.571	NaN	NaN	4.762	NaN
## 2216	104	51.92	NaN	60.000	NaN	NaN	20.000	NaN
## 2300	869	58.57	9.091	39.394	30.3030	9.091	36.364	27.2727
## 2301	220	63.18	100.000	30.000	-70.0000	0.000	10.000	10.0000
## 2302	195	52.31	NaN	30.000	NaN	NaN	20.000	NaN
## 2303	3802	56.31	17.308	37.113	19.8057	15.385	22.680	7.2958
## 2304	819	46.15	12.500	50.000	37.5000	25.000	19.231	-5.7692
## 2305	470	43.62	0.000	42.105	42.1053	0.000	15.789	15.7895
## 2306	287	60.28	100.000	45.455	-54.5455	0.000	18.182	18.1818
## 2307	376	52.93	50.000	77.778	27.7778	16.667	44.444	27.7778
## 2308	1177	60.49	10.000	41.176	31.1765	10.000	45.098	35.0980
## 2309	776	66.11	9.091	36.364	27.2727	9.091	27.273	18.1818
## 2310	614	49.35	33.333	53.333	20.0000	0.000	16.667	16.6667
## 2311	411	46.23	33.333	42.857	9.5238	0.000	14.286	14.2857
## 2312	1844	49.08	13.793	32.203	18.4103	10.345	23.729	13.3840
## 2400	264	55.30	0.000	36.364	36.3636	0.000	36.364	36.3636
## 2401	7	42.86	NaN	0.000	NaN	NaN	0.000	NaN
## 2402	670	44.78	47.368	71.429	24.0602	31.579	28.571	-3.0075
## 2403	739	47.23	26.667	53.846	27.1795	13.333	15.385	2.0513

##	2404	873	50.52	35.714	47.368	11.6541	28.571	26.316	-2.2556
##	2405	333	33.03	0.000	54.545	54.5455	20.000	18.182	-1.8182
##	2406	145	57.24	NaN	83.333	NaN	NaN	66.667	NaN
##	2500	1126	40.76	20.000	23.333	3.3333	0.000	0.000	0.0000
##	2501	88	28.41	NaN	NaN	NaN	NaN	NaN	NaN
##	2502	193	57.51	100.000	50.000	-50.0000	0.000	10.000	10.0000
##	2503	339	35.69	0.000	25.000	25.0000	20.000	50.000	30.0000
##	2504	580	40.17	0.000	11.765	11.7647	0.000	5.882	5.8824
##	2505	1022	50.29	18.519	10.000	-8.5185	22.222	10.000	-12.2222
##	2506	278	43.17	0.000	22.222	22.2222	20.000	0.000	-20.0000
##	2507	462	35.06	28.571	42.857	14.2857	14.286	42.857	28.5714
##	2508	417	40.53	0.000	33.333	33.3333	20.000	20.000	0.0000
##	2600	389	54.50	22.222	20.000	-2.2222	22.222	20.000	-2.2222
##	2601	116	49.14	50.000	0.000	-50.0000	50.000	0.000	-50.0000
##	2602	147	56.46	0.000	0.000	0.0000	0.000	0.000	0.0000
##	2603	101	45.54	NaN	0.000	NaN	NaN	0.000	NaN
##	2604	812	47.78	7.143	16.667	9.5238	7.143	25.000	17.8571
##	2605	215	45.12	0.000	50.000	50.0000	0.000	0.000	0.0000
##	2606	65	30.77	NaN	25.000	NaN	NaN	25.000	NaN
##	2607	146	63.01	0.000	0.000	0.0000	0.000	0.000	0.0000
##	2608	114	51.75	NaN	0.000	NaN	NaN	0.000	NaN
##	2609	126	72.22	0.000	0.000	0.0000	0.000	0.000	0.0000
##	2610	124	44.35	100.000	25.000	-75.0000	0.000	0.000	0.0000
##	2611	420	42.86	0.000	28.571	28.5714	0.000	42.857	42.8571
##	2612	115	52.17	0.000	0.000	0.0000	0.000	0.000	0.0000
##	2613	615	50.57	16.667	0.000	-16.6667	16.667	12.500	-4.1667
##	2614	309	55.66	20.000	0.000	-20.0000	0.000	11.111	11.1111
##	2700	4762	67.47	18.056	48.000	29.9444	25.000	40.000	15.0000
##	2701	852	59.98	33.333	64.706	31.3725	22.222	29.412	7.1895
##	2702	207	56.52	25.000	40.000	15.0000	25.000	60.000	35.0000
##	2703	432	32.41	0.000	45.455	45.4545	0.000	18.182	18.1818
##	2704	147	60.54	0.000	40.000	40.0000	50.000	0.000	-50.0000
##	2705	856	60.98	14.286	36.364	22.0779	0.000	13.636	13.6364
##	2706	262	40.08	0.000	66.667	66.6667	0.000	33.333	33.3333
##	2707	139	69.78	NaN	37.500	NaN	NaN	25.000	NaN
##	2708	221	57.01	66.667	44.444	-22.2222	66.667	44.444	-22.2222
##	2709	6	100.00	NaN	NaN	NaN	NaN	NaN	NaN
##	2710	81	48.15	NaN	50.000	NaN	NaN	100.000	NaN
##	2711	219	69.41	NaN	62.500	NaN	NaN	50.000	NaN
##	2712	755	50.33	14.286	50.000	35.7143	14.286	44.444	30.1587
##	2713	346	52.02	100.000	66.667	-33.3333	0.000	55.556	55.5556
##	2714	94	68.09	NaN	14.286	NaN	NaN	28.571	NaN
##	2715	257	51.36	NaN	62.500	NaN	NaN	37.500	NaN
##	2716	239	58.16	50.000	40.000	-10.0000	33.333	20.000	-13.3333
##	2717	274	75.91	25.000	71.429	46.4286	25.000	38.095	13.0952
##	2718	196	52.04	NaN	50.000	NaN	NaN	16.667	NaN
##	2719	607	73.48	0.000	41.667	41.6667	0.000	41.667	41.6667
##	2720	198	56.57	80.000	57.143	-22.8571	0.000	28.571	28.5714
##	2721	116	56.90	NaN	50.000	NaN	NaN	0.000	NaN
##	2722	229	52.40	0.000	40.000	40.0000	0.000	60.000	60.0000

##	2723	415	52.77	14.286	33.333	19.0476	42.857	20.000	-22.8571
##	2724	497	26.96	0.000	33.333	33.3333	66.667	33.333	-33.3333
##	2725	590	52.20	12.500	56.522	44.0217	25.000	30.435	5.4348
##	2726	335	62.09	0.000	35.294	35.2941	0.000	29.412	29.4118
##	2727	172	63.37	0.000	50.000	50.0000	0.000	50.000	50.0000
##	2728	783	62.71	66.667	44.828	-21.8391	33.333	27.586	-5.7471
##	2729	671	54.10	20.000	90.476	70.4762	0.000	71.429	71.4286
##	2730	498	57.43	37.500	55.556	18.0556	12.500	44.444	31.9444
##	2731	534	70.41	25.000	33.333	8.3333	25.000	37.500	12.5000
##	2732	1194	63.23	14.286	37.500	23.2143	14.286	14.583	0.2976
##	2733	410	56.59	28.571	52.941	24.3697	28.571	52.941	24.3697
##	2734	577	64.47	13.333	50.000	36.6667	26.667	43.750	17.0833
##	2735	1043	58.49	20.000	55.814	35.8140	30.000	41.860	11.8605
##	2736	1281	72.60	7.692	67.273	59.5804	69.231	61.818	-7.4126
##	2737	557	59.78	20.000	40.909	20.9091	0.000	18.182	18.1818
##	2738	1479	70.86	24.138	64.062	39.9246	24.138	43.750	19.6121
##	2739	2371	67.86	18.750	65.517	46.7672	12.500	51.724	39.2241
##	2740	455	52.09	50.000	30.769	-19.2308	0.000	15.385	15.3846
##	2741	479	51.36	20.000	36.842	16.8421	0.000	21.053	21.0526
##	2742	437	65.22	10.000	69.231	59.2308	20.000	30.769	10.7692
##	2743	281	43.06	100.000	90.909	-9.0909	50.000	45.455	-4.5455
##	2744	3	66.67	0.000	NaN	NaN	50.000	NaN	NaN
##	2745	240	43.33	0.000	40.000	40.0000	0.000	20.000	20.0000
##	2746	1219	52.99	0.000	23.636	23.6364	33.333	20.000	-13.3333
##	2747	84	67.86	NaN	0.000	NaN	NaN	75.000	NaN
##	2748	138	56.52	0.000	25.000	25.0000	0.000	25.000	25.0000
##	2800	835	55.93	57.143	45.161	-11.9816	57.143	19.355	-37.7880
##	2801	68	66.18	NaN	75.000	NaN	NaN	0.000	NaN
##	2802	512	71.09	50.000	56.000	6.0000	35.714	28.000	-7.7143
##	2803	101	73.27	50.000	33.333	-16.6667	50.000	16.667	-33.3333
##	2804	368	53.26	0.000	58.333	58.3333	0.000	33.333	33.3333
##	2805	289	60.21	75.000	40.000	-35.0000	50.000	40.000	-10.0000
##	2806	79	68.35	NaN	100.000	NaN	NaN	0.000	NaN
##	2807	80	32.50	NaN	100.000	NaN	NaN	50.000	NaN
##	2808	346	67.05	NaN	33.333	NaN	NaN	22.222	NaN
##	2809	299	59.20	66.667	55.556	-11.1111	33.333	11.111	-22.2222
##	2900	577	65.68	100.000	83.871	-16.1290	100.000	90.323	-9.6774
##	2901	31	51.61	NaN	100.000	NaN	NaN	0.000	NaN
##	2902	117	60.68	NaN	50.000	NaN	NaN	50.000	NaN
##	2903	1	0.00	NaN	NaN	NaN	NaN	NaN	NaN
##	2904	5	60.00	NaN	NaN	NaN	NaN	NaN	NaN
##	2905	83	78.31	NaN	66.667	NaN	NaN	33.333	NaN
##	2906	29	96.55	NaN	50.000	NaN	NaN	50.000	NaN
##	2907	32	78.12	NaN	100.000	NaN	NaN	100.000	NaN
##	2908	16	50.00	NaN	NaN	NaN	NaN	NaN	NaN
##	2909	57	80.70	NaN	87.500	NaN	NaN	50.000	NaN
##	2910	138	78.99	NaN	33.333	NaN	NaN	55.556	NaN
##	2911	55	54.55	0.000	100.000	100.0000	0.000	100.000	100.0000
##	2912	43	81.40	NaN	NaN	NaN	NaN	NaN	NaN
##	2913	47	76.60	NaN	100.000	NaN	NaN	100.000	NaN

##	2914	18	61.11	NaN	0.000	NaN	NaN	0.000	NaN
##	2915	0	NA	NA	NA	NA	NA	NA	NA
##	2916	874	64.99	50.000	68.182	18.1818	20.000	47.727	27.7273
##	2917	10	60.00	NaN	100.000	NaN	NaN	100.000	NaN
##	2918	0	NA	NA	NA	NA	NA	NA	NA
##	2919	44	86.36	100.000	100.000	0.0000	100.000	0.000	-100.0000
##	2920	4	100.00	NaN	100.000	NaN	NaN	100.000	NaN
##	2921	122	70.49	NaN	72.727	NaN	NaN	36.364	NaN
##	2922	21	90.48	NaN	NaN	NaN	NaN	NaN	NaN
##	2923	14	78.57	NaN	NaN	NaN	NaN	NaN	NaN
##	3000	91	74.73	100.000	33.333	-66.6667	0.000	0.000	0.0000
##	3001	7	85.71	NaN	0.000	NaN	NaN	0.000	NaN
##	3002	524	70.04	0.000	50.000	50.0000	0.000	35.714	35.7143
##	3003	532	59.02	50.000	42.308	-7.6923	25.000	42.308	17.3077
##	3004	1036	60.42	31.250	60.526	29.2763	25.000	34.211	9.2105
##	3005	447	50.78	33.333	84.615	51.2821	100.000	38.462	-61.5385
##	3100	657	39.42	NaN	0.000	NaN	NaN	10.000	NaN
##	3101	203	34.48	0.000	33.333	33.3333	0.000	0.000	0.0000
##	3102	166	46.39	0.000	25.000	25.0000	0.000	25.000	25.0000
##	3103	191	33.51	0.000	100.000	100.0000	0.000	0.000	0.0000
##	3104	1440	38.06	0.000	16.000	16.0000	0.000	8.000	8.0000
##	3105	332	36.45	0.000	30.769	30.7692	0.000	23.077	23.0769
##	3106	177	34.46	NaN	0.000	NaN	NaN	0.000	NaN
##	3107	620	36.29	12.500	13.636	1.1364	0.000	13.636	13.6364
##	3108	53	49.06	0.000	100.000	100.0000	0.000	0.000	0.0000
##	3109	174	36.21	50.000	20.000	-30.0000	0.000	0.000	0.0000
##	3110	201	34.33	0.000	33.333	33.3333	0.000	16.667	16.6667
##	3200	937	73.64	42.857	44.444	1.5873	31.429	29.630	-1.7989
##	3201	130	73.08	0.000	50.000	50.0000	100.000	25.000	-75.0000
##	3202	557	70.92	41.667	66.667	25.0000	33.333	54.167	20.8333
##	3203	488	78.07	44.444	68.182	23.7374	22.222	59.091	36.8687
##	3204	957	74.71	36.364	68.750	32.3864	27.273	64.583	37.3106
##	3205	710	75.21	28.571	51.515	22.9437	19.048	36.364	17.3160
##	3206	191	82.72	0.000	50.000	50.0000	0.000	25.000	25.0000
##	3207	737	74.90	55.556	61.111	5.5556	44.444	58.333	13.8889
##	3300	660	73.48	20.000	43.333	23.3333	20.000	46.667	26.6667
##	3301	604	81.79	16.667	62.162	45.4955	16.667	62.162	45.4955
##	3302	211	70.14	28.571	25.000	-3.5714	57.143	12.500	-44.6429
##	3303	677	72.38	18.182	54.839	36.6569	36.364	45.161	8.7977
##	3304	4067	77.43	38.182	67.299	29.1168	32.727	62.085	29.3580
##	3305	2255	70.55	29.630	52.000	22.3704	37.037	52.000	14.9630
##	3306	851	74.27	20.000	57.143	37.1429	40.000	45.238	5.2381
##	3307	202	69.31	0.000	25.000	25.0000	0.000	25.000	25.0000
##	3308	902	77.05	41.667	44.681	3.0142	41.667	40.426	-1.2411
##	3309	481	70.48	75.000	36.364	-38.6364	62.500	27.273	-35.2273
##	3310	1067	77.69	66.667	61.702	-4.9645	58.333	65.957	7.6241
##	3311	198	73.23	0.000	43.750	43.7500	0.000	12.500	12.5000
##	3312	2044	80.72	43.243	53.982	10.7391	32.432	51.327	18.8950
##	3313	302	64.24	0.000	37.500	37.5000	28.571	31.250	2.6786
##	3314	565	75.40	60.000	62.857	2.8571	70.000	62.857	-7.1429

## 3315	631	81.30	100.000	37.143	-62.8571	100.000	37.143	-62.8571
## 3316	1066	79.83	50.000	61.765	11.7647	37.500	61.765	24.2647
## 3317	154	67.53	50.000	44.444	-5.5556	50.000	55.556	5.5556
## 3318	324	82.41	83.333	89.474	6.1404	66.667	94.737	28.0702
## 3319	177	73.45	NaN	83.333	NaN	NaN	75.000	NaN
## 3320	482	81.12	NaN	27.027	NaN	NaN	32.432	NaN
## 3321	267	82.77	40.000	41.667	1.6667	0.000	41.667	41.6667
## 3322	211	71.56	100.000	45.455	-54.5455	50.000	18.182	-31.8182
## 3400	1448	14.57	0.000	34.615	34.6154	0.000	23.077	23.0769
## 3401	31	25.81	NaN	NaN	NaN	NaN	NaN	NaN
## 3402	68	22.06	0.000	100.000	100.0000	0.000	0.000	0.0000
## 3403	224	22.32	0.000	60.000	60.0000	0.000	0.000	0.0000
## 3404	60	30.00	0.000	100.000	100.0000	0.000	50.000	50.0000
## 3500	305	28.52	0.000	28.571	28.5714	0.000	14.286	14.2857
## 3501	6	33.33	NaN	NaN	NaN	NaN	NaN	NaN
## 3502	0	NA	NA	NA	NA	NA	NA	NA
## 3503	0	NA	NA	NA	NA	NA	NA	NA
## 3504	99	43.43	0.000	0.000	0.0000	0.000	50.000	50.0000
## 3505	22	50.00	NaN	50.000	NaN	NaN	0.000	NaN
## 3506	18	11.11	NaN	NaN	NaN	NaN	NaN	NaN
## 3600	60	65.00	NaN	NaN	NaN	NaN	NaN	NaN
## 3601	52	82.69	NaN	100.000	NaN	NaN	50.000	NaN
## 3602	17	76.47	NaN	66.667	NaN	NaN	66.667	NaN
## 3603	15	73.33	NaN	50.000	NaN	NaN	0.000	NaN
## 3604	24	70.83	NaN	100.000	NaN	NaN	100.000	NaN
## 3605	31	54.84	NaN	0.000	NaN	NaN	0.000	NaN
## 3606	0	NA	NA	NA	NA	NA	NA	NA
## 3607	48	47.92	NaN	0.000	NaN	NaN	0.000	NaN
## 3608	0	NA	NA	NA	NA	NA	NA	NA
## 3609	40	77.50	100.000	100.000	0.0000	100.000	100.000	0.0000
## 3610	73	79.45	NaN	42.857	NaN	NaN	57.143	NaN
## 3611	91	81.32	NaN	66.667	NaN	NaN	66.667	NaN
## 3612	959	70.39	28.571	52.381	23.8095	21.429	21.429	0.0000
## 3613	1	100.00	NaN	NaN	NaN	NaN	NaN	NaN
## 3614	82	59.76	NaN	25.000	NaN	NaN	25.000	NaN
## 3615	0	NA	NA	NA	NA	NA	NA	NA
## 3616	199	82.91	NaN	61.538	NaN	NaN	53.846	NaN

```
print(paste("Stats just for gendered aticles from 1 country 1996 -",
LastYearForCitationAnalysis))
```

```
## [1] "Stats just for gendered aticles from 1 country 1996 - 2012"
```

```
print(BasicStats2)
```

##	1CtyMFto14	Fem1All	Fem1 96	Fem1 14	Chg1	FemNA11	FemN96	FemN14
## 1000	147	21.769	16.667	0.000	-16.6667	17.007	16.667	25.000
## 1100	410	32.195	15.385	43.478	28.0936	18.780	11.538	23.188
## 1101	89	22.472	16.667	55.556	38.8889	12.360	0.000	33.333
## 1102	485	31.340	26.923	26.471	-0.4525	24.536	11.538	32.353
## 1103	833	32.533	15.625	30.435	14.8098	20.408	12.500	23.913

## 1104	1163	25.881	20.000	32.584	12.5843	21.410	18.182	26.966
## 1105	2180	31.330	25.000	38.690	13.6905	21.468	19.737	23.214
## 1106	619	37.157	35.484	53.191	17.7076	26.494	32.258	38.298
## 1107	277	13.718	0.000	18.182	18.1818	19.495	7.692	22.727
## 1108	243	29.630	23.810	42.857	19.0476	22.634	14.286	28.571
## 1109	300	36.667	20.000	52.381	32.3810	34.667	10.000	33.333
## 1110	845	29.231	17.021	34.545	17.5242	25.207	21.277	30.909
## 1111	198	19.192	12.500	22.581	10.0806	15.152	0.000	19.355
## 1200	34	35.294	NaN	50.000	NaN	32.353	NaN	25.000
## 1201	469	51.386	60.000	55.932	-4.0678	48.401	60.000	55.932
## 1202	386	34.197	21.429	50.000	28.5714	34.197	21.429	50.000
## 1203	470	51.702	81.818	56.522	-25.2964	51.702	72.727	50.000
## 1204	91	45.055	33.333	25.000	-8.3333	31.868	33.333	25.000
## 1205	22	18.182	NaN	0.000	NaN	18.182	NaN	0.000
## 1206	7	42.857	NaN	0.000	NaN	42.857	NaN	0.000
## 1207	164	47.561	57.143	50.000	-7.1429	41.463	71.429	35.000
## 1208	251	42.231	0.000	45.455	45.4545	40.637	0.000	45.455
## 1209	3	0.000	NaN	NaN	NaN	0.000	NaN	NaN
## 1210	46	32.609	0.000	28.571	28.5714	28.261	0.000	14.286
## 1211	333	27.027	19.048	24.138	5.0903	27.327	19.048	24.138
## 1212	118	29.661	NaN	33.333	NaN	28.814	NaN	33.333
## 1213	147	50.340	NaN	58.333	NaN	53.741	NaN	58.333
## 1300	365	30.137	33.333	42.623	9.2896	20.000	0.000	26.230
## 1301	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
## 1302	23	47.826	0.000	20.000	20.0000	43.478	50.000	20.000
## 1303	1052	34.791	27.778	44.444	16.6667	23.099	12.500	35.556
## 1304	222	36.486	33.333	35.000	1.6667	17.568	13.333	20.000
## 1305	382	45.812	35.294	51.515	16.2210	21.728	5.882	18.182
## 1306	134	37.313	33.333	20.000	-13.3333	35.075	33.333	60.000
## 1307	368	49.185	34.783	48.148	13.3655	30.163	34.783	37.037
## 1308	280	38.571	28.571	45.238	16.6667	25.000	28.571	21.429
## 1309	144	50.000	50.000	37.500	-12.5000	31.944	66.667	37.500
## 1310	273	40.293	13.333	50.000	36.6667	21.612	13.333	15.000
## 1311	650	40.769	32.558	50.980	18.4223	21.385	16.279	25.490
## 1312	767	39.505	30.612	51.786	21.1735	21.382	18.367	32.143
## 1313	187	40.642	0.000	56.250	56.2500	27.273	0.000	37.500
## 1314	450	36.889	17.857	36.842	18.9850	18.667	10.714	34.211
## 1315	88	21.591	25.000	50.000	25.0000	17.045	12.500	33.333
## 1400	257	34.630	45.455	61.905	16.4502	36.576	45.455	42.857
## 1401	129	40.310	28.571	70.000	41.4286	41.860	28.571	60.000
## 1402	271	32.841	25.000	42.308	17.3077	30.627	8.333	34.615
## 1403	419	38.425	44.444	52.941	8.4967	36.516	44.444	44.118
## 1404	56	32.143	0.000	100.000	100.0000	28.571	0.000	100.000
## 1405	301	31.229	21.569	58.333	36.7647	28.904	21.569	50.000
## 1406	278	31.655	0.000	30.000	30.0000	30.576	0.000	35.000
## 1407	282	46.809	41.667	56.522	14.8551	45.035	41.667	47.826
## 1408	494	39.879	28.000	45.714	17.7143	37.045	36.000	51.429
## 1409	372	41.398	0.000	42.424	42.4242	38.441	18.182	24.242
## 1410	47	44.681	100.000	75.000	-25.0000	40.426	100.000	25.000
## 1500	218	14.679	0.000	25.000	25.0000	19.266	12.500	18.750

## 1501	20	50.000	NaN	NaN	NaN	10.000	NaN	NaN
## 1502	78	30.769	33.333	40.000	6.6667	10.256	0.000	20.000
## 1503	121	27.273	14.286	40.000	25.7143	26.446	14.286	40.000
## 1504	12	8.333	NaN	NaN	NaN	8.333	NaN	NaN
## 1505	39	23.077	NaN	0.000	NaN	12.821	NaN	33.333
## 1506	11	18.182	NaN	0.000	NaN	18.182	NaN	0.000
## 1507	17	0.000	NaN	0.000	NaN	5.882	NaN	0.000
## 1508	23	43.478	50.000	0.000	-50.0000	34.783	0.000	100.000
## 1600	641	20.905	12.821	27.586	14.7657	15.445	10.256	17.241
## 1601	20	45.000	NaN	50.000	NaN	40.000	NaN	50.000
## 1602	243	33.745	33.333	41.379	8.0460	20.165	16.667	27.586
## 1603	45	40.000	NaN	100.000	NaN	31.111	NaN	50.000
## 1604	322	18.323	22.727	31.250	8.5227	18.323	18.182	37.500
## 1605	598	32.609	30.000	50.000	20.0000	26.923	13.333	33.333
## 1606	556	19.604	21.053	25.000	3.9474	14.568	13.158	17.857
## 1607	117	23.077	0.000	57.143	57.1429	16.239	0.000	28.571
## 1700	135	22.222	20.000	50.000	30.0000	13.333	10.000	25.000
## 1701	14	7.143	NaN	NaN	NaN	0.000	NaN	NaN
## 1702	108	25.000	20.000	30.769	10.7692	25.926	20.000	38.462
## 1703	79	8.861	0.000	0.000	0.0000	6.329	0.000	0.000
## 1704	23	8.696	0.000	NaN	NaN	8.696	0.000	NaN
## 1705	101	15.842	0.000	50.000	50.0000	13.861	0.000	25.000
## 1706	296	21.622	60.000	18.750	-41.2500	21.959	40.000	25.000
## 1707	41	7.317	NaN	0.000	NaN	2.439	NaN	0.000
## 1708	40	20.000	NaN	80.000	NaN	7.500	NaN	0.000
## 1709	49	32.653	0.000	16.667	16.6667	20.408	0.000	0.000
## 1710	147	30.612	66.667	46.667	-20.0000	22.449	33.333	20.000
## 1711	31	9.677	50.000	0.000	-50.0000	9.677	0.000	0.000
## 1712	167	14.970	10.000	50.000	40.0000	8.982	10.000	8.333
## 1800	50	28.000	NaN	100.000	NaN	18.000	NaN	100.000
## 1801	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
## 1802	96	33.333	66.667	0.000	-66.6667	21.875	33.333	0.000
## 1803	123	21.951	33.333	50.000	16.6667	19.512	33.333	0.000
## 1804	96	10.417	0.000	18.182	18.1818	9.375	0.000	36.364
## 1900	280	30.357	21.429	36.842	15.4135	28.571	21.429	31.579
## 1901	334	18.563	12.500	26.667	14.1667	13.772	6.250	13.333
## 1902	189	14.815	0.000	28.571	28.5714	10.582	0.000	28.571
## 1903	30	10.000	0.000	25.000	25.0000	13.333	0.000	25.000
## 1904	260	24.615	21.429	28.000	6.5714	23.846	14.286	28.000
## 1905	27	33.333	NaN	50.000	NaN	3.704	NaN	0.000
## 1906	282	18.085	0.000	33.333	33.3333	12.766	0.000	22.222
## 1907	405	20.494	19.048	14.815	-4.2328	17.284	4.762	29.630
## 1908	434	19.355	8.333	36.000	27.6667	13.134	4.167	16.000
## 1909	169	8.876	0.000	21.429	21.4286	11.243	0.000	7.143
## 1910	344	20.930	12.500	21.875	9.3750	20.349	12.500	21.875
## 1911	114	19.298	10.000	50.000	40.0000	25.439	10.000	25.000
## 1912	121	14.050	0.000	30.000	30.0000	11.570	0.000	20.000
## 1913	28	32.143	NaN	0.000	NaN	14.286	NaN	0.000
## 2000	184	17.935	14.286	25.000	10.7143	17.391	14.286	18.750
## 2001	72	34.722	0.000	50.000	50.0000	33.333	0.000	33.333

## 2002	630	20.000	15.000	30.909	15.9091	21.270	15.000	29.091
## 2003	302	21.192	18.182	16.667	-1.5152	18.543	18.182	11.111
## 2100	57	21.053	NaN	40.000	NaN	28.070	NaN	40.000
## 2101	2	0.000	NaN	NaN	NaN	0.000	NaN	NaN
## 2102	57	7.018	0.000	0.000	0.0000	7.018	0.000	0.000
## 2103	33	24.242	0.000	33.333	33.3333	3.030	0.000	0.000
## 2104	9	0.000	0.000	NaN	NaN	0.000	0.000	NaN
## 2105	114	24.561	0.000	45.455	45.4545	18.421	0.000	27.273
## 2200	118	23.729	23.077	16.667	-6.4103	18.644	7.692	0.000
## 2201	45	15.556	NaN	0.000	NaN	11.111	NaN	0.000
## 2202	43	11.628	NaN	0.000	NaN	23.256	NaN	50.000
## 2203	23	26.087	NaN	NaN	NaN	39.130	NaN	NaN
## 2204	104	22.115	NaN	12.500	NaN	11.538	NaN	12.500
## 2205	269	8.550	0.000	15.385	15.3846	8.550	0.000	11.538
## 2206	10	10.000	0.000	0.000	0.0000	0.000	0.000	0.000
## 2207	59	6.780	0.000	16.667	16.6667	5.085	0.000	0.000
## 2208	293	10.239	0.000	21.053	21.0526	7.167	0.000	15.789
## 2209	130	21.538	NaN	0.000	NaN	21.538	NaN	50.000
## 2210	182	9.890	0.000	16.667	16.6667	8.791	0.000	20.833
## 2211	103	14.563	NaN	12.500	NaN	7.767	NaN	16.667
## 2212	53	11.321	0.000	0.000	0.0000	7.547	0.000	33.333
## 2213	117	29.915	50.000	0.000	-50.0000	23.932	25.000	75.000
## 2214	47	19.149	0.000	100.000	100.0000	12.766	0.000	0.000
## 2215	80	5.000	NaN	0.000	NaN	7.500	NaN	0.000
## 2216	22	31.818	NaN	NaN	NaN	36.364	NaN	NaN
## 2300	344	22.384	9.091	26.087	16.9960	17.442	9.091	26.087
## 2301	85	38.824	100.000	30.000	-70.0000	30.588	0.000	40.000
## 2302	50	24.000	NaN	20.000	NaN	28.000	NaN	40.000
## 2303	1453	28.355	17.308	37.705	20.3972	21.748	15.385	24.590
## 2304	236	26.271	12.500	26.667	14.1667	16.102	25.000	13.333
## 2305	107	20.561	0.000	17.647	17.6471	17.757	0.000	11.765
## 2306	90	25.556	100.000	23.077	-76.9231	17.778	0.000	23.077
## 2307	115	38.261	50.000	64.286	14.2857	33.043	16.667	21.429
## 2308	445	29.438	10.000	37.097	27.0968	26.067	10.000	35.484
## 2309	328	32.012	9.091	34.146	25.0554	23.476	9.091	36.585
## 2310	141	29.078	33.333	31.818	-1.5152	20.567	0.000	13.636
## 2311	82	25.610	33.333	44.444	11.1111	18.293	0.000	22.222
## 2312	689	19.594	13.793	30.952	17.1593	19.448	10.345	21.429
## 2400	77	27.273	0.000	83.333	83.3333	18.182	0.000	16.667
## 2401	1	0.000	NaN	0.000	NaN	0.000	NaN	0.000
## 2402	214	38.785	47.368	50.000	2.6316	19.159	31.579	37.500
## 2403	235	48.085	26.667	61.111	34.4444	26.809	13.333	50.000
## 2404	308	42.532	35.714	57.143	21.4286	21.429	28.571	42.857
## 2405	61	40.984	0.000	57.143	57.1429	9.836	20.000	0.000
## 2406	44	52.273	NaN	50.000	NaN	25.000	NaN	16.667
## 2500	278	19.424	20.000	16.667	-3.3333	15.108	0.000	16.667
## 2501	24	4.167	NaN	0.000	NaN	8.333	NaN	0.000
## 2502	58	24.138	100.000	9.091	-90.9091	15.517	0.000	9.091
## 2503	83	21.687	0.000	20.000	20.0000	16.867	20.000	20.000
## 2504	148	15.541	0.000	30.000	30.0000	15.541	0.000	10.000

## 2505	372	19.355	18.519	20.000	1.4815	12.903	22.222	20.000
## 2506	75	20.000	0.000	22.222	22.2222	17.333	20.000	11.111
## 2507	101	27.723	28.571	0.000	-28.5714	5.941	14.286	0.000
## 2508	104	18.269	0.000	42.857	42.8571	19.231	20.000	14.286
## 2600	120	12.500	22.222	20.000	-2.2222	13.333	22.222	10.000
## 2601	43	30.233	50.000	16.667	-33.3333	20.930	50.000	16.667
## 2602	53	5.660	0.000	20.000	20.0000	1.887	0.000	0.000
## 2603	26	15.385	NaN	50.000	NaN	19.231	NaN	0.000
## 2604	270	14.074	7.143	8.696	1.5528	10.741	7.143	13.043
## 2605	66	9.091	0.000	0.000	0.0000	9.091	0.000	0.000
## 2606	9	33.333	NaN	20.000	NaN	22.222	NaN	40.000
## 2607	68	8.824	0.000	0.000	0.0000	4.412	0.000	0.000
## 2608	34	11.765	NaN	0.000	NaN	11.765	NaN	0.000
## 2609	68	7.353	0.000	9.091	9.0909	7.353	0.000	9.091
## 2610	37	5.405	100.000	0.000	-100.0000	5.405	0.000	0.000
## 2611	115	20.000	0.000	11.111	11.1111	10.435	0.000	22.222
## 2612	40	7.500	0.000	0.000	0.0000	12.500	0.000	16.667
## 2613	215	12.093	16.667	7.143	-9.5238	9.767	16.667	14.286
## 2614	123	12.195	20.000	16.667	-3.3333	9.756	0.000	8.333
## 2700	2134	39.972	18.056	44.860	26.8043	29.803	25.000	30.841
## 2701	316	56.329	33.333	65.789	32.4561	50.000	22.222	44.737
## 2702	72	45.833	25.000	50.000	25.0000	25.000	25.000	33.333
## 2703	87	22.989	0.000	63.636	63.6364	14.943	0.000	9.091
## 2704	65	24.615	0.000	28.571	28.5714	15.385	50.000	0.000
## 2705	334	23.054	14.286	20.588	6.3025	12.575	0.000	11.765
## 2706	77	29.870	0.000	28.571	28.5714	12.987	0.000	0.000
## 2707	53	49.057	NaN	25.000	NaN	30.189	NaN	25.000
## 2708	76	35.526	66.667	42.857	-23.8095	22.368	66.667	14.286
## 2709	3	100.000	NaN	NaN	NaN	100.000	NaN	NaN
## 2710	29	55.172	NaN	66.667	NaN	34.483	NaN	33.333
## 2711	96	33.333	NaN	42.857	NaN	26.042	NaN	28.571
## 2712	245	37.959	14.286	34.615	20.3297	23.265	14.286	19.231
## 2713	114	38.596	100.000	50.000	-50.0000	26.316	0.000	40.000
## 2714	39	53.846	NaN	33.333	NaN	38.462	NaN	0.000
## 2715	79	27.848	NaN	44.444	NaN	24.051	NaN	33.333
## 2716	94	41.489	50.000	66.667	16.6667	23.404	33.333	50.000
## 2717	100	64.000	25.000	61.538	36.5385	58.000	25.000	57.692
## 2718	71	47.887	NaN	58.333	NaN	38.028	NaN	41.667
## 2719	241	55.187	0.000	51.852	51.8519	49.793	0.000	51.852
## 2720	73	36.986	80.000	0.000	-80.0000	23.288	0.000	0.000
## 2721	44	25.000	NaN	75.000	NaN	18.182	NaN	25.000
## 2722	82	29.268	0.000	50.000	50.0000	14.634	0.000	16.667
## 2723	132	46.970	14.286	46.154	31.8681	29.545	42.857	30.769
## 2724	82	42.683	0.000	66.667	66.6667	30.488	66.667	16.667
## 2725	168	41.667	12.500	50.000	37.5000	22.024	25.000	25.000
## 2726	129	43.411	0.000	50.000	50.0000	18.605	0.000	18.750
## 2727	79	18.987	0.000	10.000	10.0000	13.924	0.000	10.000
## 2728	302	46.689	66.667	64.286	-2.3810	32.450	33.333	28.571
## 2729	247	57.895	20.000	75.000	55.0000	37.652	0.000	60.000
## 2730	157	44.586	37.500	38.095	0.5952	29.299	12.500	38.095

## 2731	261	45.977	25.000	39.130	14.1304	16.475	25.000	34.783
## 2732	438	27.626	14.286	28.070	13.7845	24.886	14.286	29.825
## 2733	138	39.130	28.571	40.000	11.4286	27.536	28.571	30.000
## 2734	245	37.551	13.333	71.429	58.0952	28.163	26.667	42.857
## 2735	390	48.462	20.000	52.500	32.5000	36.667	30.000	45.000
## 2736	388	59.794	7.692	67.857	60.1648	53.093	69.231	61.905
## 2737	210	37.619	20.000	47.619	27.6190	22.857	0.000	42.857
## 2738	675	51.556	24.138	65.672	41.5337	43.556	24.138	46.269
## 2739	923	51.138	18.750	56.452	37.7016	42.362	12.500	52.419
## 2740	144	34.028	50.000	25.000	-25.0000	13.889	0.000	25.000
## 2741	156	28.205	20.000	50.000	30.0000	19.231	0.000	15.000
## 2742	186	51.613	10.000	72.000	62.0000	50.538	20.000	60.000
## 2743	86	56.977	100.000	83.333	-16.6667	32.558	50.000	16.667
## 2744	2	0.000	0.000	NaN	NaN	50.000	50.000	NaN
## 2745	60	43.333	0.000	83.333	83.3333	15.000	0.000	33.333
## 2746	405	18.025	0.000	20.000	20.0000	12.346	33.333	14.286
## 2747	46	23.913	NaN	33.333	NaN	23.913	NaN	33.333
## 2748	38	15.789	0.000	25.000	25.0000	7.895	0.000	0.000
## 2800	307	47.231	57.143	47.368	-9.7744	20.847	57.143	15.789
## 2801	21	33.333	NaN	0.000	NaN	28.571	NaN	0.000
## 2802	233	42.060	50.000	66.667	16.6667	21.459	35.714	33.333
## 2803	44	47.727	50.000	100.000	50.0000	36.364	50.000	33.333
## 2804	114	47.368	0.000	50.000	50.0000	21.053	0.000	41.667
## 2805	108	51.852	75.000	61.111	-13.8889	27.778	50.000	22.222
## 2806	30	66.667	NaN	0.000	NaN	53.333	NaN	33.333
## 2807	13	76.923	NaN	100.000	NaN	7.692	NaN	0.000
## 2808	119	38.655	NaN	76.471	NaN	27.731	NaN	35.294
## 2809	120	48.333	66.667	12.500	-54.1667	15.000	33.333	25.000
## 2900	247	74.089	100.000	92.000	-8.0000	66.802	100.000	88.000
## 2901	7	71.429	NaN	100.000	NaN	42.857	NaN	50.000
## 2902	47	68.085	NaN	80.000	NaN	61.702	NaN	80.000
## 2903	0	NaN	NaN	NaN	NaN	NaN	NaN	NaN
## 2904	1	100.000	NaN	NaN	NaN	100.000	NaN	NaN
## 2905	34	61.765	NaN	57.143	NaN	58.824	NaN	71.429
## 2906	15	93.333	NaN	100.000	NaN	100.000	NaN	100.000
## 2907	14	42.857	NaN	NaN	NaN	57.143	NaN	NaN
## 2908	2	50.000	NaN	NaN	NaN	50.000	NaN	NaN
## 2909	13	46.154	NaN	0.000	NaN	23.077	NaN	50.000
## 2910	52	48.077	NaN	50.000	NaN	48.077	NaN	58.333
## 2911	25	56.000	0.000	100.000	100.0000	44.000	0.000	100.000
## 2912	20	80.000	NaN	50.000	NaN	75.000	NaN	50.000
## 2913	20	90.000	NaN	100.000	NaN	85.000	NaN	50.000
## 2914	7	57.143	NaN	50.000	NaN	42.857	NaN	100.000
## 2915	0	NA	NA	NA	NA	NA	NA	NA
## 2916	322	64.286	50.000	69.767	19.7674	45.963	20.000	44.186
## 2917	2	100.000	NaN	NaN	NaN	100.000	NaN	NaN
## 2918	0	NA	NA	NA	NA	NA	NA	NA
## 2919	25	92.000	100.000	100.000	0.0000	84.000	100.000	100.000
## 2920	2	0.000	NaN	NaN	NaN	0.000	NaN	NaN
## 2921	53	49.057	NaN	100.000	NaN	35.849	NaN	75.000

## 2922	15	73.333	NaN	33.333	NaN	80.000	NaN	33.333
## 2923	7	14.286	NaN	0.000	NaN	42.857	NaN	25.000
## 3000	53	66.038	100.000	100.000	0.0000	58.491	0.000	100.000
## 3001	3	33.333	NaN	100.000	NaN	0.000	NaN	0.000
## 3002	266	31.579	0.000	42.857	42.8571	21.805	0.000	28.571
## 3003	207	51.691	50.000	57.143	7.1429	35.749	25.000	38.095
## 3004	417	50.120	31.250	58.974	27.7244	35.731	25.000	41.026
## 3005	150	41.333	33.333	58.824	25.4902	36.000	100.000	41.176
## 3100	187	14.439	NaN	10.000	NaN	3.743	NaN	0.000
## 3101	38	10.526	0.000	0.000	0.0000	7.895	0.000	25.000
## 3102	52	7.692	0.000	0.000	0.0000	7.692	0.000	0.000
## 3103	50	16.000	0.000	0.000	0.0000	14.000	0.000	33.333
## 3104	380	16.579	0.000	27.586	27.5862	12.105	0.000	6.897
## 3105	77	5.195	0.000	0.000	0.0000	5.195	0.000	0.000
## 3106	46	26.087	NaN	50.000	NaN	8.696	NaN	12.500
## 3107	145	13.103	12.500	25.000	12.5000	8.966	0.000	16.667
## 3108	18	11.111	0.000	100.000	100.0000	11.111	0.000	100.000
## 3109	44	9.091	50.000	25.000	-25.0000	4.545	0.000	0.000
## 3110	51	41.176	0.000	100.000	100.0000	27.451	0.000	0.000
## 3200	476	48.950	42.857	52.174	9.3168	37.395	31.429	43.478
## 3201	58	41.379	0.000	85.714	85.7143	37.931	100.000	42.857
## 3202	228	42.982	41.667	44.000	2.3333	44.737	33.333	52.000
## 3203	241	59.751	44.444	74.074	29.6296	37.344	22.222	40.741
## 3204	415	61.446	36.364	55.814	19.4503	56.145	27.273	55.814
## 3205	354	40.678	28.571	47.826	19.2547	30.508	19.048	39.130
## 3206	115	43.478	0.000	28.571	28.5714	30.435	0.000	57.143
## 3207	323	48.297	55.556	61.538	5.9829	49.226	44.444	66.667
## 3300	244	47.131	20.000	43.478	23.4783	45.492	20.000	39.130
## 3301	279	60.215	16.667	65.789	49.1228	58.423	16.667	65.789
## 3302	104	45.192	28.571	60.000	31.4286	38.462	57.143	40.000
## 3303	335	39.104	18.182	40.909	22.7273	37.612	36.364	36.364
## 3304	1781	52.330	38.182	57.407	19.2256	51.319	32.727	55.556
## 3305	1042	37.908	29.630	34.000	4.3704	35.893	37.037	33.000
## 3306	335	58.209	20.000	57.143	37.1429	53.433	40.000	52.381
## 3307	96	29.167	0.000	33.333	33.3333	14.583	0.000	0.000
## 3308	364	35.714	41.667	60.000	18.3333	33.791	41.667	53.333
## 3309	213	42.723	75.000	42.857	-32.1429	41.784	62.500	47.619
## 3310	507	50.888	66.667	55.556	-11.1111	51.085	58.333	50.000
## 3311	57	43.860	0.000	50.000	50.0000	40.351	0.000	50.000
## 3312	940	43.617	43.243	46.875	3.6318	40.957	32.432	43.750
## 3313	122	28.689	0.000	44.444	44.4444	22.951	28.571	22.222
## 3314	241	58.506	60.000	62.500	2.5000	56.432	70.000	65.625
## 3315	282	41.844	100.000	37.778	-62.2222	42.908	100.000	37.778
## 3316	419	45.107	50.000	51.471	1.4706	44.630	37.500	55.882
## 3317	70	60.000	50.000	45.455	-4.5455	57.143	50.000	45.455
## 3318	174	82.759	83.333	81.250	-2.0833	80.460	66.667	81.250
## 3319	62	64.516	NaN	83.333	NaN	64.516	NaN	100.000
## 3320	194	25.773	NaN	36.364	NaN	27.835	NaN	31.818
## 3321	143	44.755	40.000	20.000	-20.0000	39.860	0.000	40.000
## 3322	83	46.988	100.000	62.500	-37.5000	38.554	50.000	62.500

## 3400	114	30.702	0.000	57.143	57.1429	28.947	0.000	28.571
## 3401	6	50.000	NaN	NaN	NaN	0.000	NaN	NaN
## 3402	4	0.000	0.000	NaN	NaN	25.000	0.000	NaN
## 3403	24	41.667	0.000	100.000	100.0000	25.000	0.000	0.000
## 3404	11	9.091	0.000	NaN	NaN	36.364	0.000	NaN
## 3500	63	41.270	0.000	28.571	28.5714	33.333	0.000	28.571
## 3501	1	0.000	NaN	NaN	NaN	0.000	NaN	NaN
## 3502	0	NA	NA	NA	NA	NA	NA	NA
## 3503	0	NA	NA	NA	NA	NA	NA	NA
## 3504	29	13.793	0.000	0.000	0.0000	13.793	0.000	0.000
## 3505	3	33.333	NaN	NaN	NaN	0.000	NaN	NaN
## 3506	1	0.000	NaN	NaN	NaN	0.000	NaN	NaN
## 3600	33	57.576	NaN	50.000	NaN	51.515	NaN	50.000
## 3601	24	45.833	NaN	40.000	NaN	37.500	NaN	40.000
## 3602	8	25.000	NaN	0.000	NaN	62.500	NaN	100.000
## 3603	7	42.857	NaN	NaN	NaN	14.286	NaN	NaN
## 3604	10	80.000	NaN	100.000	NaN	80.000	NaN	100.000
## 3605	11	54.545	NaN	0.000	NaN	27.273	NaN	0.000
## 3606	0	NA	NA	NA	NA	NA	NA	NA
## 3607	13	30.769	NaN	NaN	NaN	15.385	NaN	NaN
## 3608	0	NA	NA	NA	NA	NA	NA	NA
## 3609	20	75.000	100.000	33.333	-66.6667	90.000	100.000	33.333
## 3610	31	35.484	NaN	60.000	NaN	6.452	NaN	20.000
## 3611	46	65.217	NaN	66.667	NaN	60.870	NaN	66.667
## 3612	395	37.975	28.571	35.294	6.7227	31.646	21.429	37.255
## 3613	1	0.000	NaN	NaN	NaN	0.000	NaN	NaN
## 3614	27	33.333	NaN	50.000	NaN	25.926	NaN	0.000
## 3615	0	NA	NA	NA	NA	NA	NA	NA
## 3616	77	67.532	NaN	66.667	NaN	71.429	NaN	58.333
##	ChgN 1CtyMF96 1CtyMF14							
## 1000	8.3333	6	4					
## 1100	11.6499	26	69					
## 1101	33.3333	6	9					
## 1102	20.8145	26	34					
## 1103	11.4130	32	46					
## 1104	8.7845	55	89					
## 1105	3.4774	76	168					
## 1106	6.0398	31	47					
## 1107	15.0350	13	44					
## 1108	14.2857	21	7					
## 1109	23.3333	10	21					
## 1110	9.6325	47	55					
## 1111	19.3548	8	31					
## 1200	NaN	0	4					
## 1201	-4.0678	10	59					
## 1202	28.5714	14	44					
## 1203	-22.7273	11	46					
## 1204	-8.3333	6	4					
## 1205	NaN	0	1					
## 1206	NaN	0	1					

## 1207	-36.4286	7	20
## 1208	45.4545	3	22
## 1209	NaN	0	0
## 1210	14.2857	2	7
## 1211	5.0903	21	29
## 1212	NaN	0	12
## 1213	NaN	0	24
## 1300	26.2295	15	61
## 1301	NaN	0	0
## 1302	-30.0000	2	5
## 1303	23.0556	72	90
## 1304	6.6667	15	20
## 1305	12.2995	17	33
## 1306	26.6667	6	10
## 1307	2.2544	23	27
## 1308	-7.1429	7	42
## 1309	-29.1667	6	8
## 1310	1.6667	15	20
## 1311	9.2111	43	51
## 1312	13.7755	49	56
## 1313	37.5000	5	32
## 1314	23.4962	28	38
## 1315	20.8333	8	6
## 1400	-2.5974	11	21
## 1401	31.4286	7	10
## 1402	26.2821	12	26
## 1403	-0.3268	9	34
## 1404	100.0000	1	5
## 1405	28.4314	51	12
## 1406	35.0000	9	20
## 1407	6.1594	12	23
## 1408	15.4286	25	35
## 1409	6.0606	11	33
## 1410	-75.0000	1	4
## 1500	6.2500	16	16
## 1501	NaN	0	0
## 1502	20.0000	6	5
## 1503	25.7143	7	5
## 1504	NaN	0	0
## 1505	NaN	0	3
## 1506	NaN	0	2
## 1507	NaN	0	3
## 1508	100.0000	2	1
## 1600	6.9850	39	29
## 1601	NaN	0	2
## 1602	10.9195	12	29
## 1603	NaN	0	2
## 1604	19.3182	22	16
## 1605	20.0000	30	48
## 1606	4.6992	38	28

## 1607	28.5714	3	7
## 1700	15.0000	10	4
## 1701	NaN	0	0
## 1702	18.4615	5	13
## 1703	0.0000	4	5
## 1704	NaN	1	0
## 1705	25.0000	2	4
## 1706	-15.0000	5	32
## 1707	NaN	0	4
## 1708	NaN	0	5
## 1709	0.0000	2	6
## 1710	-13.3333	6	15
## 1711	0.0000	2	3
## 1712	-1.6667	10	12
## 1800	NaN	0	1
## 1801	NaN	0	0
## 1802	-33.3333	3	1
## 1803	-33.3333	6	2
## 1804	36.3636	3	11
## 1900	10.1504	14	19
## 1901	7.0833	16	15
## 1902	28.5714	13	7
## 1903	25.0000	2	4
## 1904	13.7143	14	25
## 1905	NaN	0	2
## 1906	22.2222	15	27
## 1907	24.8677	21	27
## 1908	11.8333	24	25
## 1909	7.1429	2	14
## 1910	9.3750	16	32
## 1911	15.0000	10	4
## 1912	20.0000	8	10
## 1913	NaN	0	1
## 2000	4.4643	7	16
## 2001	33.3333	2	6
## 2002	14.0909	20	55
## 2003	-7.0707	11	18
## 2100	NaN	0	5
## 2101	NaN	0	0
## 2102	0.0000	3	4
## 2103	0.0000	3	3
## 2104	NaN	1	0
## 2105	27.2727	6	11
## 2200	-7.6923	13	6
## 2201	NaN	0	4
## 2202	NaN	0	2
## 2203	NaN	0	0
## 2204	NaN	0	16
## 2205	11.5385	5	26
## 2206	0.0000	1	1

## 2207	0.0000	2	6
## 2208	15.7895	11	19
## 2209	NaN	0	6
## 2210	20.8333	4	24
## 2211	NaN	0	24
## 2212	33.3333	1	3
## 2213	50.0000	4	4
## 2214	0.0000	4	1
## 2215	NaN	0	11
## 2216	NaN	0	0
## 2300	16.9960	11	23
## 2301	40.0000	1	10
## 2302	NaN	0	5
## 2303	9.2055	52	122
## 2304	-11.6667	8	15
## 2305	11.7647	5	17
## 2306	23.0769	1	13
## 2307	4.7619	6	14
## 2308	25.4839	10	62
## 2309	27.4945	11	41
## 2310	13.6364	3	22
## 2311	22.2222	3	9
## 2312	11.0837	29	42
## 2400	16.6667	1	6
## 2401	NaN	0	1
## 2402	5.9211	19	16
## 2403	36.6667	15	18
## 2404	14.2857	14	28
## 2405	-20.0000	5	7
## 2406	NaN	0	6
## 2500	16.6667	5	36
## 2501	NaN	0	1
## 2502	9.0909	1	11
## 2503	0.0000	5	10
## 2504	10.0000	1	10
## 2505	-2.2222	27	20
## 2506	-8.8889	5	9
## 2507	-14.2857	7	5
## 2508	-5.7143	5	7
## 2600	-12.2222	9	10
## 2601	-33.3333	2	6
## 2602	0.0000	1	5
## 2603	NaN	0	2
## 2604	5.9006	14	23
## 2605	0.0000	3	5
## 2606	NaN	0	5
## 2607	0.0000	4	3
## 2608	NaN	0	3
## 2609	9.0909	1	11
## 2610	0.0000	1	3

## 2611	22.2222	5	9
## 2612	16.6667	2	6
## 2613	-2.3810	12	14
## 2614	8.3333	5	12
## 2700	5.8411	72	214
## 2701	22.5146	9	38
## 2702	8.3333	4	6
## 2703	9.0909	1	11
## 2704	-50.0000	2	7
## 2705	11.7647	7	34
## 2706	0.0000	2	7
## 2707	NaN	0	4
## 2708	-52.3810	3	7
## 2709	NaN	0	0
## 2710	NaN	0	3
## 2711	NaN	0	14
## 2712	4.9451	7	26
## 2713	40.0000	1	10
## 2714	NaN	0	3
## 2715	NaN	0	9
## 2716	16.6667	6	6
## 2717	32.6923	4	26
## 2718	NaN	0	12
## 2719	51.8519	5	27
## 2720	0.0000	5	1
## 2721	NaN	0	4
## 2722	16.6667	5	6
## 2723	-12.0879	7	13
## 2724	-50.0000	3	6
## 2725	0.0000	8	24
## 2726	18.7500	1	16
## 2727	10.0000	3	10
## 2728	-4.7619	9	28
## 2729	60.0000	5	20
## 2730	25.5952	8	21
## 2731	9.7826	4	23
## 2732	15.5388	7	57
## 2733	1.4286	7	10
## 2734	16.1905	15	21
## 2735	15.0000	10	40
## 2736	-7.3260	13	84
## 2737	42.8571	5	21
## 2738	22.1307	29	67
## 2739	39.9194	16	124
## 2740	25.0000	4	16
## 2741	15.0000	5	20
## 2742	40.0000	10	25
## 2743	-33.3333	2	6
## 2744	NaN	2	0
## 2745	33.3333	1	6

## 2746	-19.0476	3	35
## 2747	NaN	0	6
## 2748	0.0000	2	4
## 2800	-41.3534	14	19
## 2801	NaN	0	1
## 2802	-2.3810	14	12
## 2803	-16.6667	4	3
## 2804	41.6667	2	12
## 2805	-27.7778	4	18
## 2806	NaN	0	3
## 2807	NaN	0	1
## 2808	NaN	0	17
## 2809	-8.3333	6	8
## 2900	-12.0000	4	25
## 2901	NaN	0	2
## 2902	NaN	0	5
## 2903	NaN	0	0
## 2904	NaN	0	0
## 2905	NaN	0	7
## 2906	NaN	0	1
## 2907	NaN	0	0
## 2908	NaN	0	0
## 2909	NaN	0	2
## 2910	NaN	0	12
## 2911	100.0000	2	1
## 2912	NaN	0	2
## 2913	NaN	0	2
## 2914	NaN	0	2
## 2915	NA	NA	NA
## 2916	24.1860	10	43
## 2917	NaN	0	0
## 2918	NA	NA	NA
## 2919	0.0000	1	4
## 2920	NaN	0	0
## 2921	NaN	0	4
## 2922	NaN	0	3
## 2923	NaN	0	4
## 3000	100.0000	1	1
## 3001	NaN	0	1
## 3002	28.5714	8	28
## 3003	13.0952	4	21
## 3004	16.0256	16	39
## 3005	-58.8235	3	17
## 3100	NaN	0	10
## 3101	25.0000	3	4
## 3102	0.0000	1	5
## 3103	33.3333	2	3
## 3104	6.8966	6	29
## 3105	0.0000	4	3
## 3106	NaN	0	8

## 3107	16.6667	8	12
## 3108	100.0000	1	1
## 3109	0.0000	2	4
## 3110	0.0000	1	1
## 3200	12.0497	35	23
## 3201	-57.1429	1	7
## 3202	18.6667	12	25
## 3203	18.5185	9	27
## 3204	28.5412	11	43
## 3205	20.0828	21	23
## 3206	57.1429	2	7
## 3207	22.2222	9	39
## 3300	19.1304	5	23
## 3301	49.1228	12	38
## 3302	-17.1429	7	5
## 3303	0.0000	11	22
## 3304	22.8283	55	216
## 3305	-4.0370	27	100
## 3306	12.3810	5	42
## 3307	0.0000	2	9
## 3308	11.6667	12	30
## 3309	-14.8810	8	21
## 3310	-8.3333	12	54
## 3311	50.0000	1	8
## 3312	11.3176	37	96
## 3313	-6.3492	7	9
## 3314	-4.3750	10	32
## 3315	-62.2222	2	45
## 3316	18.3824	8	68
## 3317	-4.5455	2	11
## 3318	14.5833	6	16
## 3319	NaN	0	6
## 3320	NaN	0	22
## 3321	40.0000	5	10
## 3322	12.5000	2	8
## 3400	28.5714	3	7
## 3401	NaN	0	0
## 3402	NaN	1	0
## 3403	0.0000	1	1
## 3404	NaN	2	0
## 3500	28.5714	2	7
## 3501	NaN	0	0
## 3502	NA	NA	NA
## 3503	NA	NA	NA
## 3504	0.0000	1	2
## 3505	NaN	0	0
## 3506	NaN	0	0
## 3600	NaN	0	2
## 3601	NaN	0	5
## 3602	NaN	0	1

```
## 3603      NaN      0      0
## 3604      NaN      0      1
## 3605      NaN      0      1
## 3606      NA      NA      NA
## 3607      NaN      0      0
## 3608      NA      NA      NA
## 3609 -66.6667      1      3
## 3610      NaN      0      5
## 3611      NaN      0      3
## 3612  15.8263     14     51
## 3613      NaN      0      0
## 3614      NaN      0      2
## 3615      NA      NA      NA
## 3616      NaN      0     12
```

```
print("GenderTeamSize in 2018")
```

```
## [1] "GenderTeamSize in 2018"
```

```
print(GenderTeamSize)
```

```
##      Articles FirstF FirstM   FirstP LastF LastM   LastP
## 1000         43  4.991  3.222 0.1561888  3.915  3.693  1.00000
## 1100         61  4.324  3.167 0.0412065  4.473  3.550  0.09936
## 1101          8  2.466  3.104 1.0000000  2.080  3.438  0.21880
## 1102         77  4.375  4.197 0.7105899  4.245  4.271  0.91180
## 1103         87  3.273  3.149 0.8423549  3.257  3.171  0.88684
## 1104         69  3.113  3.126 0.9010913  2.374  3.367  0.05410
## 1105        138  3.147  3.126 0.7100579  3.061  3.167  0.66561
## 1106         62  4.489  3.965 0.5160729  4.459  4.089  0.57710
## 1107         15  3.873  3.563 0.8626843  4.949  3.210  0.08392
## 1108         31  4.112  4.768 0.5035723  4.371  4.495  0.69896
## 1109         44  4.040  4.242 0.8954912  4.185  4.110  0.54371
## 1110         64  3.732  3.768 0.7781727  3.461  3.904  0.29469
## 1111         29  4.291  4.268 0.6262880  4.571  4.149  0.73873
## 1200          4  1.000  2.000 0.6373519  2.000  1.000  0.63735
## 1201         43  1.698  2.222 0.2242700  1.878  1.931  0.79597
## 1202         60  1.393  1.021 0.0007104  1.263  1.123  0.21871
## 1203         37  1.374  1.256 0.6957664  1.514  1.051  0.03320
## 1204          9  3.464  1.219 0.0557168  2.000  1.488  0.49003
## 1205         NA      NA      NA      NA      NA      NA      NA
## 1206         NA      NA      NA      NA      NA      NA      NA
## 1207         24  1.000  1.173 0.1059059  1.059  1.122  0.57967
## 1208         18  1.149  1.000 0.2215751  1.000  1.149  0.22158
## 1209         NA      NA      NA      NA      NA      NA      NA
## 1210          5  1.260  1.000 0.6830914  1.000  1.260  0.68309
## 1211         39  1.631  1.055 0.0037859  1.473  1.143  0.25393
## 1212         17  1.732  1.113 0.1053484  1.442  1.194  0.56933
## 1213         31  1.315  1.134 0.4337568  1.130  1.408  0.21288
## 1300         56  4.563  3.275 0.0233597  4.425  3.771  0.22582
## 1301         NA      NA      NA      NA      NA      NA      NA
```


## 1302	6	3.438	8.000	0.2347637	2.828	4.681	0.34756
## 1303	49	4.431	4.703	0.4585165	4.656	4.512	0.79794
## 1304	18	5.061	3.413	0.0162412	5.000	4.208	0.84402
## 1305	38	2.851	3.850	0.2305776	2.116	3.973	0.02623
## 1306	8	4.579	6.491	0.6488071	5.264	6.160	1.00000
## 1307	23	4.124	4.112	0.7535895	4.043	4.178	0.87528
## 1308	18	4.280	4.394	0.8100969	3.245	4.819	0.36257
## 1309	7	3.728	3.464	0.8436689	3.420	3.834	0.71882
## 1310	20	3.945	1.904	0.1561669	4.782	2.902	0.12090
## 1311	40	3.756	5.438	0.0464458	3.846	5.056	0.24852
## 1312	44	4.107	3.754	0.8677350	4.118	3.785	0.89355
## 1313	15	5.108	4.277	0.4072154	4.460	4.743	0.57285
## 1314	34	3.923	3.627	0.5283842	2.850	4.313	0.04323
## 1315	5	3.000	2.714	0.7670969	3.464	2.466	1.00000
## 1400	11	2.080	1.622	0.5877935	1.732	1.737	1.00000
## 1401	10	2.913	2.621	1.0000000	2.711	2.749	0.46946
## 1402	27	1.587	2.007	0.2147674	1.762	1.914	0.71245
## 1403	31	2.312	1.836	0.2616127	2.280	1.872	0.47673
## 1404	10	1.817	1.768	1.0000000	2.182	1.316	0.20941
## 1405	18	2.213	1.556	0.2707061	2.352	1.479	0.11866
## 1406	28	2.177	1.736	0.2535713	2.008	1.912	0.90357
## 1407	18	2.289	1.463	0.1009264	2.140	1.513	0.21397
## 1408	44	1.995	2.090	0.6348785	2.080	2.017	0.91128
## 1409	46	1.690	1.991	0.2570753	1.810	1.883	0.72664
## 1410	4	3.000	1.000	0.2482131	NA	NA	NA
## 1500	19	4.120	3.433	0.9180177	3.464	3.580	1.00000
## 1501	NA	NA	NA	NA	NA	NA	NA
## 1502	9	5.485	3.957	0.5132919	5.000	4.344	0.69489
## 1503	12	4.899	3.467	0.8262696	6.928	3.235	0.09973
## 1504	NA	NA	NA	NA	NA	NA	NA
## 1505	NA	NA	NA	NA	NA	NA	NA
## 1506	NA	NA	NA	NA	NA	NA	NA
## 1507	5	3.000	2.913	0.6926328	NA	NA	NA
## 1508	5	3.000	3.224	1.0000000	NA	NA	NA
## 1600	36	3.855	3.236	0.8720446	4.752	3.177	0.18952
## 1601	5	2.236	4.000	1.0000000	1.000	4.642	0.12815
## 1602	16	3.807	2.747	0.1559631	1.817	3.341	0.11549
## 1603	NA	NA	NA	NA	NA	NA	NA
## 1604	9	7.862	2.994	0.0262394	3.663	4.547	0.32106
## 1605	20	4.243	3.638	0.5084319	4.190	3.706	0.52266
## 1606	24	3.961	3.269	0.4402614	3.795	3.328	0.52014
## 1607	4	4.000	2.000	0.6373519	NA	NA	NA
## 1700	9	5.000	2.294	0.1642985	5.000	2.294	0.16430
## 1701	NA	NA	NA	NA	NA	NA	NA
## 1702	4	1.000	2.289	0.3457786	1.414	2.449	0.41422
## 1703	NA	NA	NA	NA	NA	NA	NA
## 1704	NA	NA	NA	NA	NA	NA	NA
## 1705	5	4.000	1.732	0.2635525	NA	NA	NA
## 1706	15	2.000	2.266	0.7924601	2.000	2.266	0.79246
## 1707	NA	NA	NA	NA	NA	NA	NA

## 1708	NA	NA	NA	NA	NA	NA	NA
## 1709	5	2.000	2.520	1.0000000	1.000	2.828	0.26355
## 1710	13	3.224	2.217	0.0902802	2.449	2.495	1.00000
## 1711	5	3.000	3.130	1.0000000	NA	NA	NA
## 1712	11	4.000	2.169	0.3245335	NA	NA	NA
## 1800	3	3.000	3.000	NaN	3.000	3.000	NaN
## 1801	NA	NA	NA	NA	NA	NA	NA
## 1802	7	2.828	2.169	0.6848604	2.289	2.378	0.85304
## 1803	9	3.000	2.213	0.8418270	2.000	2.380	0.45070
## 1804	NA	NA	NA	NA	NA	NA	NA
## 1900	20	3.022	2.440	0.4344840	2.823	2.509	0.90437
## 1901	27	3.113	2.104	0.0278123	2.449	2.514	0.74743
## 1902	12	4.899	2.625	0.0405926	3.000	2.905	1.00000
## 1903	NA	NA	NA	NA	NA	NA	NA
## 1904	30	2.380	3.031	0.3191562	2.773	2.893	0.83808
## 1905	NA	NA	NA	NA	NA	NA	NA
## 1906	20	3.623	3.363	0.6966740	3.360	3.529	0.79956
## 1907	42	3.596	2.459	0.0095076	2.913	2.980	0.98229
## 1908	40	4.032	2.841	0.0422890	3.631	3.183	0.55430
## 1909	25	3.764	2.127	0.0082909	2.000	2.579	0.47514
## 1910	16	1.763	1.931	0.8195635	1.414	2.204	0.18965
## 1911	9	1.414	1.575	1.0000000	1.587	1.513	0.66501
## 1912	5	2.828	1.000	0.2635525	4.000	2.000	0.45606
## 1913	3	3.162	6.000	0.6666667	NA	NA	NA
## 2000	20	4.000	1.687	0.0975125	2.402	1.543	0.07990
## 2001	10	1.000	1.963	0.2684813	1.861	1.817	1.00000
## 2002	58	2.146	1.836	0.3208739	2.111	1.839	0.42664
## 2003	38	1.866	1.761	0.8614700	1.919	1.737	0.65955
## 2100	6	3.476	2.520	0.6428348	3.000	2.952	1.00000
## 2101	NA	NA	NA	NA	NA	NA	NA
## 2102	8	1.414	2.245	0.1904303	1.000	2.208	0.08648
## 2103	5	1.414	2.520	0.3329216	1.000	2.378	0.23568
## 2104	4	1.414	2.000	0.6170751	1.000	2.000	0.24821
## 2105	28	2.717	2.093	0.1816685	2.363	2.242	0.48148
## 2200	9	2.000	1.861	0.8367590	2.000	1.861	0.83676
## 2201	3	2.000	1.732	1.0000000	NA	NA	NA
## 2202	NA	NA	NA	NA	NA	NA	NA
## 2203	2	1.000	4.000	1.0000000	1.000	4.000	1.00000
## 2204	18	4.644	4.523	0.5481792	6.160	4.180	0.06680
## 2205	28	3.064	2.684	0.6178237	2.000	2.860	0.71250
## 2206	NA	NA	NA	NA	NA	NA	NA
## 2207	NA	NA	NA	NA	NA	NA	NA
## 2208	8	6.000	2.208	0.1742314	4.899	2.000	0.11956
## 2209	16	5.092	2.101	0.0360317	2.289	2.527	0.62470
## 2210	23	3.722	2.808	0.2569006	2.000	3.061	0.12723
## 2211	22	4.095	3.347	0.4203351	2.000	3.706	0.05259
## 2212	5	3.000	3.310	1.0000000	NA	NA	NA
## 2213	18	4.227	3.416	0.3636134	3.663	3.614	0.87042
## 2214	NA	NA	NA	NA	NA	NA	NA
## 2215	21	3.706	3.018	0.3326517	2.000	3.276	0.23060

## 2216	5	3.107	1.414	0.2361370	2.000	2.340	1.00000
## 2300	33	2.245	2.343	0.8770269	2.426	2.237	0.51676
## 2301	10	2.289	2.354	1.0000000	2.000	2.375	0.85654
## 2302	10	4.481	2.826	0.2016024	4.243	3.035	0.50605
## 2303	97	3.373	3.370	0.8072876	3.210	3.420	0.63802
## 2304	26	3.130	2.521	0.1970875	2.352	2.930	0.27529
## 2305	19	3.464	2.754	0.3732753	2.884	3.062	0.77409
## 2306	11	2.048	3.420	0.3074796	1.000	3.380	0.04174
## 2307	18	4.430	5.091	0.5142903	4.860	4.349	0.46690
## 2308	51	2.534	2.828	0.5951905	2.604	2.788	0.83047
## 2309	33	2.161	2.866	0.1522486	2.189	2.754	0.30587
## 2310	30	3.505	4.034	0.4181558	3.245	3.851	0.58788
## 2311	21	3.608	3.281	0.8001034	3.302	3.437	0.87803
## 2312	59	3.815	3.309	0.2506785	2.933	3.648	0.07858
## 2400	11	4.091	3.536	0.6295030	4.527	3.337	0.24694
## 2401	NA	NA	NA	NA	NA	NA	NA
## 2402	7	4.678	2.000	0.1998489	4.472	3.390	0.66914
## 2403	13	6.642	5.201	0.8860918	4.899	6.143	0.69221
## 2404	19	4.093	3.532	1.0000000	4.618	3.529	0.54340
## 2405	11	3.915	3.129	0.5069490	3.464	3.552	0.80663
## 2406	6	5.573	3.000	0.3657123	6.506	3.000	0.09520
## 2500	30	3.337	2.900	0.6312520	NA	NA	NA
## 2501	NA	NA	NA	NA	NA	NA	NA
## 2502	10	6.297	5.024	0.1326223	5.000	5.698	1.00000
## 2503	4	5.000	4.610	1.0000000	3.162	7.000	0.22067
## 2504	17	4.583	3.168	0.4493105	3.000	3.329	0.83586
## 2505	10	5.000	4.610	1.0000000	5.000	4.610	1.00000
## 2506	9	5.000	5.135	1.0000000	NA	NA	NA
## 2507	7	5.828	3.364	0.3590121	2.884	5.701	0.09873
## 2508	15	3.764	3.957	0.7558445	3.000	4.153	0.21279
## 2600	20	1.682	1.168	0.0593464	1.682	1.168	0.05935
## 2601	NA	NA	NA	NA	NA	NA	NA
## 2602	NA	NA	NA	NA	NA	NA	NA
## 2603	NA	NA	NA	NA	NA	NA	NA
## 2604	12	1.414	1.931	0.5739658	2.000	1.781	0.84653
## 2605	2	2.000	1.000	1.0000000	NA	NA	NA
## 2606	4	3.000	2.621	1.0000000	3.000	2.621	1.00000
## 2607	NA	NA	NA	NA	NA	NA	NA
## 2608	NA	NA	NA	NA	NA	NA	NA
## 2609	NA	NA	NA	NA	NA	NA	NA
## 2610	4	2.000	1.000	0.2482131	NA	NA	NA
## 2611	7	2.449	2.702	0.8343194	2.289	2.913	0.33973
## 2612	NA	NA	NA	NA	NA	NA	NA
## 2613	NA	NA	NA	NA	NA	NA	NA
## 2614	NA	NA	NA	NA	NA	NA	NA
## 2700	100	3.870	3.468	0.3120324	3.349	3.875	0.46395
## 2701	34	4.633	4.040	0.9128319	4.282	4.471	0.47895
## 2702	5	1.414	1.442	1.0000000	1.817	1.000	0.33292
## 2703	11	4.618	2.904	0.2222811	4.583	3.396	0.62781
## 2704	5	5.000	5.429	1.0000000	NA	NA	NA

## 2705	22	5.709	5.069	0.8359712	5.241	5.301	1.00000
## 2706	6	4.356	2.828	0.4811200	4.899	3.310	0.63859
## 2707	8	3.780	3.314	0.7614225	2.000	4.189	0.06191
## 2708	9	2.449	1.888	0.6104918	2.213	2.048	1.00000
## 2709	NA	NA	NA	NA	NA	NA	NA
## 2710	2	2.000	3.000	1.0000000	NA	NA	NA
## 2711	8	2.667	4.380	0.2662364	2.449	4.213	0.28172
## 2712	18	5.961	3.751	0.4212108	5.746	4.046	0.47189
## 2713	9	5.440	4.481	0.4287943	4.644	5.733	0.80242
## 2714	7	1.000	2.349	0.5811695	2.646	1.888	1.00000
## 2715	8	5.502	6.542	0.5437005	6.257	5.650	0.87934
## 2716	5	5.292	4.718	1.0000000	4.000	5.207	0.71680
## 2717	21	3.293	4.036	0.2538217	2.952	3.869	0.20014
## 2718	6	6.910	3.634	0.0765225	5.000	5.014	1.00000
## 2719	36	2.944	2.526	0.7425432	3.191	2.385	0.29319
## 2720	7	4.262	4.932	0.7212767	3.464	5.055	0.84508
## 2721	2	9.000	4.000	1.0000000	NA	NA	NA
## 2722	5	2.449	2.466	1.0000000	3.107	1.732	0.55362
## 2723	15	8.617	4.447	0.2190184	7.114	5.209	0.77204
## 2724	15	4.129	3.826	0.8986690	4.573	3.635	0.22637
## 2725	23	5.180	4.370	0.4684918	6.321	4.270	0.22153
## 2726	17	5.309	4.404	0.5078026	5.547	4.393	0.52142
## 2727	2	6.000	3.000	1.0000000	6.000	3.000	1.00000
## 2728	29	4.201	3.726	0.6249797	3.431	4.142	0.45829
## 2729	21	3.376	3.162	0.9027304	3.112	4.050	0.34062
## 2730	18	3.833	4.327	0.7506528	4.224	3.907	0.85591
## 2731	24	4.860	2.803	0.0436810	4.099	2.993	0.30423
## 2732	48	4.318	3.714	0.2599356	3.513	4.006	0.55848
## 2733	17	3.231	4.801	0.1816982	3.231	4.801	0.18170
## 2734	16	2.671	1.668	0.1150630	2.560	1.817	0.29806
## 2735	43	5.124	4.282	0.6365284	4.285	5.085	0.10102
## 2736	110	1.623	1.690	0.8685799	1.526	1.856	0.18937
## 2737	22	4.376	3.895	0.8655964	4.091	4.083	0.96558
## 2738	64	2.799	2.869	0.7492462	2.539	3.067	0.23734
## 2739	116	3.748	3.894	0.3984804	3.694	3.912	0.40276
## 2740	13	5.045	4.392	1.0000000	7.348	4.207	0.22714
## 2741	19	3.994	2.798	0.2096904	3.310	3.159	0.95918
## 2742	13	4.475	5.477	0.3791818	3.464	5.486	0.02043
## 2743	11	3.728	7.000	0.1500155	3.438	4.430	0.26781
## 2744	NA	NA	NA	NA	NA	NA	NA
## 2745	5	3.873	5.130	0.7609067	9.000	3.873	0.26355
## 2746	55	4.553	4.181	0.6153234	3.815	4.387	0.23189
## 2747	NA	NA	NA	NA	NA	NA	NA
## 2748	4	6.000	4.718	1.0000000	6.000	4.718	1.00000
## 2800	31	4.556	3.794	0.2639485	4.107	4.124	0.97959
## 2801	4	6.082	1.000	0.3457786	NA	NA	NA
## 2802	25	3.347	1.946	0.0384693	3.061	2.487	0.53628
## 2803	6	2.646	2.378	1.0000000	1.000	2.952	0.36571
## 2804	12	4.450	5.727	0.1907450	3.557	5.828	0.10418
## 2805	10	1.682	2.720	0.3159124	1.968	2.449	0.82364

## 2806	NA	NA	NA	NA	NA	NA	NA
## 2807	NA	NA	NA	NA	NA	NA	NA
## 2808	18	4.545	3.688	0.7388659	5.318	3.633	0.19499
## 2809	9	4.441	2.115	0.4406860	4.000	3.105	1.00000
## 2900	31	2.275	4.373	0.1457930	2.407	4.000	0.08781
## 2901	NA	NA	NA	NA	NA	NA	NA
## 2902	4	2.449	3.873	0.4142162	2.449	3.873	0.41422
## 2903	NA	NA	NA	NA	NA	NA	NA
## 2904	NA	NA	NA	NA	NA	NA	NA
## 2905	3	3.464	6.000	0.6666667	4.000	4.243	1.00000
## 2906	2	3.000	3.000	NaN	3.000	3.000	NaN
## 2907	NA	NA	NA	NA	NA	NA	NA
## 2908	NA	NA	NA	NA	NA	NA	NA
## 2909	8	2.583	3.000	0.8136637	2.632	2.632	1.00000
## 2910	9	2.520	1.414	0.3123214	2.639	1.000	0.03998
## 2911	NA	NA	NA	NA	NA	NA	NA
## 2912	NA	NA	NA	NA	NA	NA	NA
## 2913	NA	NA	NA	NA	NA	NA	NA
## 2914	NA	NA	NA	NA	NA	NA	NA
## 2915	NA	NA	NA	NA	NA	NA	NA
## 2916	44	5.713	3.809	0.0223268	4.966	5.073	0.82942
## 2917	NA	NA	NA	NA	NA	NA	NA
## 2918	NA	NA	NA	NA	NA	NA	NA
## 2919	NA	NA	NA	NA	NA	NA	NA
## 2920	NA	NA	NA	NA	NA	NA	NA
## 2921	11	2.446	2.080	0.8333272	2.213	2.416	0.92240
## 2922	NA	NA	NA	NA	NA	NA	NA
## 2923	NA	NA	NA	NA	NA	NA	NA
## 3000	3	4.000	3.464	1.0000000	NA	NA	NA
## 3001	NA	NA	NA	NA	NA	NA	NA
## 3002	14	5.159	4.116	0.1887209	4.043	4.955	1.00000
## 3003	26	4.106	3.480	0.3446996	3.523	3.894	0.79295
## 3004	38	3.217	3.527	0.6729096	3.450	3.279	0.47488
## 3005	13	4.078	1.732	0.1088246	5.335	2.783	0.01739
## 3100	NA	NA	NA	NA	NA	NA	NA
## 3101	6	2.000	1.861	1.0000000	NA	NA	NA
## 3102	4	3.000	1.587	0.3457786	2.000	1.817	1.00000
## 3103	NA	NA	NA	NA	NA	NA	NA
## 3104	25	4.229	3.235	0.7574739	5.477	3.238	0.10588
## 3105	13	2.711	3.520	0.6936098	2.884	3.366	1.00000
## 3106	NA	NA	NA	NA	NA	NA	NA
## 3107	22	4.932	2.299	0.0275966	4.932	2.299	0.02760
## 3108	NA	NA	NA	NA	NA	NA	NA
## 3109	5	5.000	1.682	0.2765005	NA	NA	NA
## 3110	6	3.464	4.949	0.1526614	3.000	4.743	0.22780
## 3200	27	3.117	2.126	0.4339600	3.177	2.286	0.33639
## 3201	4	3.464	2.449	0.4142162	4.000	2.621	0.34578
## 3202	24	2.221	2.087	0.9490461	2.072	2.304	0.60731
## 3203	22	2.928	3.326	0.4728868	3.086	2.996	0.78563
## 3204	48	2.324	2.476	0.6901901	2.300	2.504	0.73223

## 3205	33	2.677	1.977	0.1599152	2.702	2.114	0.35645
## 3206	4	2.000	2.000	1.0000000	1.000	2.520	0.61708
## 3207	36	2.022	2.398	0.3668928	2.044	2.334	0.50867
## 3300	30	1.947	1.899	0.9476992	1.872	1.962	0.82806
## 3301	37	1.728	1.842	0.7646374	1.835	1.669	0.75198
## 3302	8	3.464	1.260	0.0857117	2.000	1.575	0.61708
## 3303	31	1.745	1.842	0.7218325	1.731	1.838	0.78543
## 3304	211	1.871	1.646	0.1224513	1.827	1.742	0.58070
## 3305	100	1.731	1.754	0.8001018	1.842	1.640	0.44747
## 3306	42	2.704	2.177	0.4422050	2.539	2.405	0.82575
## 3307	4	6.000	1.817	0.5000000	3.000	2.289	1.00000
## 3308	47	1.544	1.118	0.0293348	1.368	1.242	0.35706
## 3309	11	1.565	1.575	1.0000000	1.442	1.622	0.82569
## 3310	47	1.478	1.240	0.2821839	1.541	1.118	0.04382
## 3311	16	3.326	3.067	0.9137844	3.162	3.180	1.00000
## 3312	113	1.630	1.397	0.1172767	1.617	1.420	0.27167
## 3313	16	2.305	3.557	0.3438231	2.091	3.574	0.22194
## 3314	35	1.567	1.445	0.6334599	1.574	1.434	0.70284
## 3315	35	1.211	1.294	0.6201501	1.211	1.294	0.62015
## 3316	68	1.357	1.374	0.9759720	1.374	1.347	0.96396
## 3317	9	1.000	1.149	0.5023350	1.149	1.000	0.50233
## 3318	19	1.362	2.828	0.0726369	1.414	3.000	0.18143
## 3319	12	2.595	2.646	1.0000000	2.934	1.817	0.29793
## 3320	37	1.149	1.080	1.0000000	1.260	1.028	0.05780
## 3321	12	2.297	1.919	0.6079827	2.297	1.919	0.60798
## 3322	11	1.741	1.570	0.6233205	1.414	1.702	0.79978
## 3400	26	3.687	3.722	1.0000000	4.141	3.589	0.77709
## 3401	NA	NA	NA	NA	NA	NA	NA
## 3402	NA	NA	NA	NA	NA	NA	NA
## 3403	5	5.646	2.828	0.1386406	NA	NA	NA
## 3404	NA	NA	NA	NA	NA	NA	NA
## 3500	7	6.000	5.144	0.2800872	5.000	5.440	0.57700
## 3501	NA	NA	NA	NA	NA	NA	NA
## 3502	NA	NA	NA	NA	NA	NA	NA
## 3503	NA	NA	NA	NA	NA	NA	NA
## 3504	NA	NA	NA	NA	NA	NA	NA
## 3505	2	5.000	6.000	1.0000000	NA	NA	NA
## 3506	NA	NA	NA	NA	NA	NA	NA
## 3600	NA	NA	NA	NA	NA	NA	NA
## 3601	NA	NA	NA	NA	NA	NA	NA
## 3602	3	5.477	6.000	1.0000000	6.000	5.000	0.47950
## 3603	2	3.000	1.000	1.0000000	NA	NA	NA
## 3604	NA	NA	NA	NA	NA	NA	NA
## 3605	NA	NA	NA	NA	NA	NA	NA
## 3606	NA	NA	NA	NA	NA	NA	NA
## 3607	NA	NA	NA	NA	NA	NA	NA
## 3608	NA	NA	NA	NA	NA	NA	NA
## 3609	NA	NA	NA	NA	NA	NA	NA
## 3610	7	5.241	2.913	0.0692004	4.243	3.175	0.58566
## 3611	9	3.634	3.302	1.0000000	3.888	2.884	0.50780

```
## 3612      42  3.436  3.584 0.8771653 2.244 3.960 0.01949
## 3613      NA     NA     NA          NA     NA     NA     NA
## 3614       4  2.000  2.884 0.6373519 2.000 2.884 0.63735
## 3615      NA     NA     NA          NA     NA     NA     NA
## 3616      13  2.952  2.402 0.6485822 2.945 2.493 0.71092
```

```
print(RegCoef)
```

```
##          FFA1          FLA1          2          3          4          5+
## 1000 -2.516e-01 -0.0210235  0.591067  0.787978  0.6857301  0.875030
## 1100  6.408e-02 -0.0434533  0.251214  0.194767  0.3041591  0.334053
## 1101  2.004e-01  0.0093691  0.147771  0.182336  0.6475053  0.482995
## 1102  3.995e-03 -0.0022608  0.045038  0.166245  0.2643636  0.263118
## 1103 -4.326e-02  0.0335208  0.203674  0.251678  0.3065984  0.345102
## 1104 -2.007e-02 -0.0022357  0.106239  0.091270  0.1069452  0.201046
## 1105 -1.321e-02 -0.0237666  0.160223  0.158504  0.2057400  0.253253
## 1106  1.098e-01 -0.0432076  0.345103  0.411643  0.3487532  0.464979
## 1107 -6.225e-02 -0.0985061  0.196359  0.182917  0.2657909  0.356617
## 1108 -1.007e-01  0.0019389  0.040240  0.142664  0.1952244  0.281538
## 1109  5.068e-02 -0.1139776  0.151917  0.201887  0.2602460  0.256890
## 1110 -8.946e-03 -0.0115449  0.190044  0.204824  0.3363674  0.423085
## 1111 -1.389e-01 -0.0136980  0.184217  0.099676  0.2340679  0.279667
## 1200      NA          NA          NA          NA          NA          NA
## 1201  1.191e-02  0.0493355  0.089053  0.081899  0.2708046  0.286125
## 1202 -1.393e-01  0.1787583  0.158816  0.410529  0.1889737  0.701242
## 1203  2.908e-02  0.0191976  0.264490  0.346685  0.3299680 -0.148412
## 1204  8.766e-02  0.0904159  0.478774  0.193956 -0.1702633  0.228250
## 1205      NA          NA          NA          NA          NA          NA
## 1206      NA          NA          NA          NA          NA          NA
## 1207  1.612e-01  0.1586506  0.355753  0.583171  0.8163831  0.525054
## 1208  4.505e-01 -0.4972307  0.565240  1.311675  0.4993148  0.292652
## 1209      NA          NA          NA          NA          NA          NA
## 1210      NA          NA          NA          NA          NA          NA
## 1211 -3.428e-01  0.4361947  0.441440  0.455855  1.4348871  0.228572
## 1212 -3.629e-01  0.3553593  0.359498  1.476957  2.0870081  0.041286
## 1213  3.871e-01 -0.3129892  0.430272 -0.849823 -0.2743586 -1.444795
## 1300  1.234e-01 -0.1595865  0.131621 -0.009838  0.2562646  0.279197
## 1301      NA          NA          NA          NA          NA          NA
## 1302      NA          NA          NA          NA          NA          NA
## 1303  2.338e-02 -0.0130416  0.130835  0.146361  0.1908866  0.289540
## 1304  7.565e-02 -0.1160108  0.158045  0.120140  0.1908578  0.093474
## 1305 -1.452e-02 -0.0711571  0.090268  0.134670  0.2635928  0.238742
## 1306 -3.957e-02 -0.0453731 -0.256667 -0.011119 -0.1726329 -0.011259
## 1307  4.903e-02  0.0269564  0.134365  0.176336  0.2675297  0.335541
## 1308 -8.590e-02  0.0688549  0.591704  0.561534  0.6502027  0.809435
## 1309  3.134e-02 -0.0534321 -0.031863  0.081526 -0.0064717  0.113041
## 1310 -3.189e-02  0.0444561  0.189057  0.209655  0.2384242  0.280325
## 1311  6.529e-05  0.0070832  0.122764  0.162188  0.1602405  0.207183
## 1312  7.269e-04 -0.0462142  0.174342  0.178771  0.1934580  0.223282
## 1313 -2.389e-02 -0.0685894  0.168323  0.198594  0.2953595  0.327653
```

## 1314	-4.822e-02	-0.0744368	0.091744	0.127031	0.2633677	0.309818
## 1315	1.853e-01	-0.2698955	-0.343150	-0.231768	-0.2363900	-0.102641
## 1400	2.020e-02	0.3588740	0.289703	0.381631	0.6244777	-0.101857
## 1401	-1.576e-01	0.0892529	0.129060	0.214908	-0.3559982	0.281352
## 1402	-3.906e-02	0.0863274	0.218682	0.236371	0.8165119	0.580645
## 1403	-7.643e-02	0.0407368	0.263735	0.108157	0.5578528	-0.071831
## 1404	NA	NA	NA	NA	NA	NA
## 1405	-9.336e-02	0.1557296	0.183024	0.242296	0.0831461	0.020655
## 1406	-1.570e-01	0.1424291	0.155718	-0.020568	0.4655231	0.179274
## 1407	7.073e-03	0.1644970	0.096032	0.240599	0.2349096	0.156549
## 1408	-6.244e-02	0.0431415	0.209361	0.183932	0.2061843	0.733150
## 1409	1.335e-01	0.0215628	0.194921	-0.063214	-0.1065946	-0.214464
## 1410	1.270e-02	0.0559313	0.475340	0.287863	-0.0821805	0.580735
## 1500	-2.198e-02	-0.2895222	0.322558	0.382542	0.4332197	0.554169
## 1501	NA	NA	NA	NA	NA	NA
## 1502	1.416e-01	-0.3907155	0.195091	-0.131531	0.0328795	0.043240
## 1503	-1.018e-01	-0.0837068	-0.086392	-0.006582	-0.0406235	-0.087715
## 1504	NA	NA	NA	NA	NA	NA
## 1505	NA	NA	NA	NA	NA	NA
## 1506	NA	NA	NA	NA	NA	NA
## 1507	NA	NA	NA	NA	NA	NA
## 1508	NA	NA	NA	NA	NA	NA
## 1600	5.391e-02	0.0209745	0.040489	-0.050868	0.0544665	0.199544
## 1601	NA	NA	NA	NA	NA	NA
## 1602	2.285e-02	-0.0081040	0.159795	0.258651	0.1541850	0.303013
## 1603	1.736e-01	-0.1905175	-0.365356	-0.383144	-0.1493232	-0.803613
## 1604	2.648e-02	0.0379271	0.045528	-0.098945	-0.0786830	-0.025497
## 1605	3.398e-02	-0.0152178	0.055033	0.017768	0.0532673	0.122939
## 1606	7.117e-02	0.0321582	0.131064	0.160685	0.1040971	0.158424
## 1607	1.071e-01	0.2968632	0.292188	0.287425	0.3394825	0.278083
## 1700	3.041e-01	-0.1477723	0.141238	0.272491	0.1887549	0.247156
## 1701	NA	NA	NA	NA	NA	NA
## 1702	2.519e-01	0.0997924	0.251370	0.273592	0.4117699	0.177987
## 1703	-3.278e-01	0.4048563	0.278710	0.298552	1.1081950	-0.092716
## 1704	NA	NA	NA	NA	NA	NA
## 1705	-5.148e-03	0.0910975	0.035558	-0.092917	0.1205490	-0.872315
## 1706	1.325e-01	-0.1503175	-0.062406	0.093628	0.0273987	-0.037308
## 1707	NA	NA	NA	NA	NA	NA
## 1708	NA	NA	NA	NA	NA	NA
## 1709	3.105e-01	-0.6916303	0.197051	-0.409636	1.1362934	0.087197
## 1710	1.606e-01	-0.0641871	-0.189611	-0.081005	0.1350950	-0.261630
## 1711	NA	NA	NA	NA	NA	NA
## 1712	2.718e-01	-0.2649196	0.325903	0.338332	0.6594053	0.745741
## 1800	NA	NA	NA	NA	NA	NA
## 1801	NA	NA	NA	NA	NA	NA
## 1802	-1.200e-01	0.1454815	0.154961	0.362559	-0.0432893	0.284397
## 1803	1.754e-02	0.1609150	0.105090	0.102838	-0.1426564	-0.533587
## 1804	-3.127e-02	-0.1485038	0.307076	0.821462	0.4050536	-0.003938
## 1900	1.554e-01	-0.0819933	0.298840	0.415665	0.3057295	0.546854
## 1901	-3.928e-02	-0.0046999	-0.059072	0.042374	-0.1060956	-0.057775

##	1902	-7.982e-03	-0.0625201	0.083829	-0.090936	0.0340395	0.321553
##	1903	NA	NA	NA	NA	NA	NA
##	1904	-1.738e-01	0.1450320	-0.014485	0.033798	-0.0810027	0.003922
##	1905	NA	NA	NA	NA	NA	NA
##	1906	1.004e-01	-0.0551944	0.106130	0.036151	0.0059495	0.115290
##	1907	-4.614e-02	-0.0017666	0.078858	0.070128	0.0879714	0.115322
##	1908	1.928e-02	-0.0174201	0.031828	0.057914	-0.0285465	0.152644
##	1909	9.843e-02	0.1860966	-0.042830	-0.058006	0.2467144	0.446803
##	1910	6.295e-02	0.0285877	0.120004	0.058008	0.1782046	0.184513
##	1911	9.407e-02	-0.0768126	0.122669	0.010368	0.0507732	0.275696
##	1912	2.307e-01	-0.0888537	0.319221	0.216777	0.4424906	0.283935
##	1913	NA	NA	NA	NA	NA	NA
##	2000	2.488e-01	0.3787133	0.152633	0.488511	0.8962184	0.065794
##	2001	-9.119e-01	0.7394864	0.468809	0.129424	0.9831827	0.969380
##	2002	2.227e-02	0.0274376	0.245320	0.303440	0.5780405	0.645612
##	2003	1.108e-02	-0.0152029	0.269319	0.241229	0.1435455	0.570065
##	2100	9.301e-02	0.0238768	0.815356	0.608171	0.7229133	1.143282
##	2101	NA	NA	NA	NA	NA	NA
##	2102	1.085e+00	-0.2070767	0.775311	0.003280	-0.1555921	-0.380766
##	2103	NA	NA	NA	NA	NA	NA
##	2104	NA	NA	NA	NA	NA	NA
##	2105	1.407e-01	-0.4183272	-0.049016	0.225533	-0.0085302	0.376137
##	2200	2.118e-01	-0.1922842	0.063310	-0.159391	-0.0437662	-0.304740
##	2201	-3.077e-02	-0.0273621	0.176184	0.388327	0.0832126	0.490832
##	2202	5.433e-02	-0.3975614	0.273311	0.445797	0.3949494	-0.023111
##	2203	NA	NA	NA	NA	NA	NA
##	2204	6.231e-02	-0.3708098	0.284439	0.355101	0.3516481	0.314274
##	2205	-9.884e-03	-0.0122925	0.231746	0.028988	0.5926987	0.439533
##	2206	NA	NA	NA	NA	NA	NA
##	2207	7.428e-01	-0.3965813	0.327592	0.263614	1.6864410	-0.186054
##	2208	1.645e-01	-0.4054893	0.202852	0.364635	0.3118417	-0.043077
##	2209	2.946e-01	-0.2336038	0.117596	0.143051	-0.5193703	-0.039512
##	2210	2.824e-01	-0.4848014	0.068460	0.330299	0.0213370	0.360080
##	2211	1.830e-01	-0.3530594	0.072484	-0.038068	0.0951258	0.121886
##	2212	6.218e-01	-0.3577797	-0.284170	-0.061190	0.0592509	0.410655
##	2213	-1.553e-01	-0.1065864	0.019295	0.163530	0.1724913	0.133776
##	2214	9.124e-01	-0.4249275	-0.428969	-0.723287	0.3110079	-0.337468
##	2215	2.353e-02	-0.0864070	-0.027258	-0.329393	-0.4006016	-0.055439
##	2216	NA	NA	NA	NA	NA	NA
##	2300	1.398e-01	-0.1322562	0.163586	0.335695	0.3851625	0.285021
##	2301	-3.057e-02	0.0840270	0.267808	0.040741	-0.2576973	-0.098893
##	2302	1.132e-02	0.0062712	-0.053409	0.324573	0.4653608	-0.026067
##	2303	1.688e-02	-0.0574306	0.126148	0.125421	0.1908057	0.288934
##	2304	-1.620e-03	0.0281142	0.181218	0.266972	0.1088127	0.239282
##	2305	-8.078e-02	0.0822633	0.027726	-0.088292	-0.0280025	0.049223
##	2306	-1.788e-02	-0.0158186	-0.047267	0.085562	0.3213245	0.214827
##	2307	-4.160e-02	0.1238985	0.272803	0.167664	0.1067612	0.198602
##	2308	7.185e-02	-0.0584563	0.194496	0.176672	0.1624556	0.277450
##	2309	1.226e-02	-0.0840396	0.008659	0.024637	0.1179046	0.186953
##	2310	9.219e-02	0.1111795	0.199414	0.140718	0.1439095	0.233467

##	2311	-5.671e-03	0.2698538	-0.110285	0.089260	0.0401950	0.204596
##	2312	-1.098e-01	0.0459350	0.072569	0.127110	0.1022904	0.243240
##	2400	2.014e-01	0.0732639	0.114606	-0.190315	-0.2072534	-0.119857
##	2401	NA	NA	NA	NA	NA	NA
##	2402	-3.674e-02	-0.0677658	0.218184	0.340575	0.2756124	0.276856
##	2403	5.715e-02	0.0097282	0.233753	0.015865	0.0439760	0.241302
##	2404	6.362e-04	0.0733568	0.114065	0.057584	0.0967692	0.107131
##	2405	NA	NA	NA	NA	NA	NA
##	2406	-3.532e-02	0.1182057	-0.849903	-0.652864	-0.6318799	-0.622163
##	2500	1.492e-01	0.0600670	0.024656	-0.169264	-0.0142002	0.022616
##	2501	NA	NA	NA	NA	NA	NA
##	2502	2.168e-01	-0.2969525	0.046431	-0.239533	-0.3499393	-0.213991
##	2503	5.111e-02	-0.2180737	0.080635	-0.204097	-0.0966956	-0.135503
##	2504	-2.560e-02	0.0349403	0.184884	0.155859	0.2313418	0.204972
##	2505	5.464e-02	-0.0492148	-0.007614	-0.037520	-0.0230298	-0.096647
##	2506	7.929e-02	0.0247946	-0.165306	-0.232815	-0.0865596	0.007635
##	2507	7.111e-02	0.0230956	0.104986	-0.065727	0.0829561	-0.007513
##	2508	9.043e-02	0.0534897	-0.081876	-0.057364	-0.0469726	-0.120222
##	2600	3.390e-01	0.0431519	0.173760	-0.005923	-0.2723045	1.007443
##	2601	3.153e-01	0.0431705	0.245664	-0.121260	-0.2846261	-0.343583
##	2602	3.930e-01	0.1536335	0.080286	-0.653010	-1.0346432	-0.228855
##	2603	NA	NA	NA	NA	NA	NA
##	2604	1.156e-01	-0.2108566	0.120800	0.381472	0.2666609	0.406656
##	2605	NA	NA	NA	NA	NA	NA
##	2606	NA	NA	NA	NA	NA	NA
##	2607	-1.090e-01	-0.1223729	0.283570	1.125534	-0.1643444	-0.419637
##	2608	NA	NA	NA	NA	NA	NA
##	2609	-2.632e-01	0.7651890	0.376552	0.544314	0.1843208	0.003032
##	2610	NA	NA	NA	NA	NA	NA
##	2611	3.562e-01	-0.2145573	0.124331	0.151207	-0.0273779	0.312082
##	2612	1.031e+00	-0.3913323	0.152362	0.496000	-1.1274238	-0.713470
##	2613	2.058e-01	0.1236697	0.138348	0.388675	0.5659211	0.312495
##	2614	1.376e-01	-0.2485661	0.075759	0.257606	0.1452749	0.387454
##	2700	6.079e-02	-0.0460949	0.561576	0.692566	0.7592345	0.907082
##	2701	3.094e-02	0.0513392	0.354928	0.417863	0.4623676	0.386562
##	2702	9.512e-02	0.0916195	-0.014841	-0.209077	-0.0267241	0.140764
##	2703	1.912e-01	0.0726033	0.267650	0.603617	0.2419219	0.572266
##	2704	1.267e-01	-0.1409404	0.102710	0.216642	0.1848178	0.652527
##	2705	1.768e-02	0.1187781	-0.075161	0.220617	0.1858202	0.337842
##	2706	3.053e-01	0.2329446	-0.143816	0.554832	0.3073234	1.142680
##	2707	-2.067e-02	0.0415996	0.194509	0.215451	0.3000705	0.291464
##	2708	1.087e-01	0.0708378	0.237478	0.496255	0.2308112	0.144223
##	2709	NA	NA	NA	NA	NA	NA
##	2710	NA	NA	NA	NA	NA	NA
##	2711	-1.188e-01	-0.0121752	0.265230	0.311937	0.4499848	0.220505
##	2712	-1.091e-01	0.1036325	0.098278	0.172449	0.1995553	0.279658
##	2713	-8.764e-02	-0.0955725	0.468579	0.334228	0.5262199	0.458128
##	2714	NA	NA	NA	NA	NA	NA
##	2715	8.295e-02	-0.2501197	-0.337778	-0.457633	-0.0015667	-0.002328
##	2716	1.365e-01	0.1497378	0.238494	0.052857	0.1440434	0.095021

##	2717	-3.881e-03	0.0544440	0.105022	0.299917	0.7022527	0.552846
##	2718	-7.226e-02	0.2708641	0.254221	0.356938	0.4464900	0.447999
##	2719	1.053e-01	-0.0069285	0.265430	0.190642	0.4159976	0.315557
##	2720	1.071e-01	-0.0489385	-0.005465	0.188490	0.0922575	0.198218
##	2721	2.329e-01	0.0140099	-0.455843	-0.549640	-0.5796723	-0.352975
##	2722	1.420e-01	-0.0716922	0.097125	-0.048599	-0.1780243	-0.006774
##	2723	-3.640e-02	-0.0786531	-0.021292	0.019833	0.1743750	0.379849
##	2724	-1.990e-02	0.1742320	-0.189374	-0.032327	0.1397955	0.170166
##	2725	-3.344e-02	0.0101613	0.136311	0.075702	0.2157619	0.213206
##	2726	-4.666e-02	-0.0929877	0.340224	0.255155	0.3398362	0.348806
##	2727	NA	NA	NA	NA	NA	NA
##	2728	8.637e-02	0.0612181	0.082080	0.186190	0.1255584	0.362624
##	2729	7.786e-02	-0.0455303	0.291924	0.299498	0.3877787	0.486988
##	2730	2.083e-02	0.0535350	-0.030500	0.127899	0.0922712	0.238379
##	2731	2.341e-03	-0.0976267	0.338334	0.486020	0.4818408	0.600282
##	2732	-4.165e-02	-0.0155194	0.182231	0.282457	0.2488681	0.360810
##	2733	-1.316e-01	0.1965703	0.542506	0.181717	0.4967131	0.564678
##	2734	1.698e-01	-0.0876603	0.040317	0.015304	0.0124832	0.001754
##	2735	-8.316e-02	0.0675410	0.246884	0.328036	0.4922811	0.491739
##	2736	3.886e-02	-0.0476145	0.018404	0.172785	0.2675051	0.271824
##	2737	-1.574e-03	0.0030781	0.003308	0.049173	0.1507503	0.197047
##	2738	8.482e-02	0.0377672	0.216468	0.276042	0.3467449	0.374537
##	2739	-4.261e-02	-0.0125179	0.416453	0.450405	0.5120209	0.501986
##	2740	1.241e-01	0.1249195	-0.082179	-0.003024	0.0278446	0.177828
##	2741	-1.279e-02	0.0307287	0.176803	0.188318	0.3335587	0.502558
##	2742	-2.405e-01	-0.1749817	0.103202	0.226391	0.1319735	0.415629
##	2743	5.992e-02	0.0780195	0.376452	0.553222	0.6249708	0.572915
##	2744	NA	NA	NA	NA	NA	NA
##	2745	4.327e-01	0.3627840	0.509112	-0.001440	0.1542147	0.298186
##	2746	-4.789e-02	0.0257397	-0.094580	-0.082947	0.0325108	0.216696
##	2747	4.435e-02	0.1918652	-1.147365	-1.070905	-0.7279992	-0.753309
##	2748	NA	NA	NA	NA	NA	NA
##	2800	2.218e-02	-0.0735887	0.393441	0.281143	0.2782737	0.356650
##	2801	NA	NA	NA	NA	NA	NA
##	2802	-3.585e-02	-0.0165580	0.056175	0.178555	0.1246208	0.263181
##	2803	6.586e-02	0.0058172	-0.101659	-0.122501	0.0214259	-0.126164
##	2804	6.856e-02	0.0770660	0.471442	0.492648	0.5809776	0.537866
##	2805	1.446e-02	-0.2325987	0.007187	0.359445	0.5042362	0.455598
##	2806	NA	NA	NA	NA	NA	NA
##	2807	NA	NA	NA	NA	NA	NA
##	2808	1.576e-01	0.0851115	0.291702	0.341657	0.4213671	0.551600
##	2809	-1.926e-01	-0.0802304	0.384360	0.236292	0.4321622	0.464741
##	2900	-2.668e-01	-0.1149453	0.009432	0.246078	0.5320083	0.483336
##	2901	NA	NA	NA	NA	NA	NA
##	2902	-4.193e-02	0.0403282	0.100335	-0.313973	0.6089702	0.217259
##	2903	NA	NA	NA	NA	NA	NA
##	2904	NA	NA	NA	NA	NA	NA
##	2905	NA	NA	NA	NA	NA	NA
##	2906	NA	NA	NA	NA	NA	NA
##	2907	NA	NA	NA	NA	NA	NA

## 2908	NA	NA	NA	NA	NA	NA
## 2909	NA	NA	NA	NA	NA	NA
## 2910	4.205e-01	-0.3090737	0.049889	-0.885725	0.9530444	-0.239242
## 2911	NA	NA	NA	NA	NA	NA
## 2912	NA	NA	NA	NA	NA	NA
## 2913	NA	NA	NA	NA	NA	NA
## 2914	NA	NA	NA	NA	NA	NA
## 2915	NA	NA	NA	NA	NA	NA
## 2916	-3.013e-02	-0.0105455	0.176048	0.338801	0.4341016	0.342728
## 2917	NA	NA	NA	NA	NA	NA
## 2918	NA	NA	NA	NA	NA	NA
## 2919	NA	NA	NA	NA	NA	NA
## 2920	NA	NA	NA	NA	NA	NA
## 2921	3.294e-01	0.2800204	0.212770	0.289466	0.7496950	0.378683
## 2922	NA	NA	NA	NA	NA	NA
## 2923	NA	NA	NA	NA	NA	NA
## 3000	-2.998e-01	0.0892011	0.018894	0.364966	1.0266850	0.602227
## 3001	NA	NA	NA	NA	NA	NA
## 3002	5.117e-02	0.0629390	-0.254426	-0.149969	-0.0944173	0.088135
## 3003	-4.453e-02	0.0576646	0.119654	0.472983	0.5128804	0.572199
## 3004	5.216e-02	-0.0425989	0.296449	0.310283	0.4280047	0.484585
## 3005	6.962e-02	0.1985388	0.039719	-0.158991	-0.1723928	0.024612
## 3100	-1.636e-02	-0.0302331	0.346196	0.189979	0.2288525	0.092071
## 3101	NA	NA	NA	NA	NA	NA
## 3102	NA	NA	NA	NA	NA	NA
## 3103	-8.684e-02	0.0614634	0.461637	0.227018	-1.0284739	0.111165
## 3104	2.091e-01	-0.0348786	-0.189194	-0.302291	-0.2215366	-0.133940
## 3105	4.913e-01	-0.5847468	-0.141880	-0.179360	-0.3993421	-0.214315
## 3106	1.280e-01	-0.0270235	-0.498841	-0.170920	0.0007102	-0.323291
## 3107	-2.082e-01	0.0866849	0.273262	0.229366	0.1905686	0.115821
## 3108	NA	NA	NA	NA	NA	NA
## 3109	-7.709e-01	0.6877437	-0.335876	0.035681	-0.0191467	0.045172
## 3110	1.578e-01	0.4754492	0.011753	0.234994	0.0535610	0.055534
## 3200	-3.418e-02	-0.0421751	0.270763	0.315377	0.3354187	0.270508
## 3201	3.896e-01	-0.1082818	0.176338	0.318654	-0.3348136	-0.086416
## 3202	-2.361e-02	0.0326684	0.200128	0.157187	0.5161764	0.335409
## 3203	-1.068e-01	-0.0883907	0.279033	0.279449	0.4010420	0.439180
## 3204	-4.734e-02	-0.0053154	0.172726	0.221609	0.3361156	0.333561
## 3205	-8.419e-02	0.0688494	0.008368	0.073831	0.1154676	0.214483
## 3206	7.378e-02	-0.0427833	0.205213	0.541017	0.3856747	0.338177
## 3207	-7.391e-02	-0.0429204	0.126847	0.258821	0.1068642	0.122459
## 3300	8.169e-03	-0.0252647	0.019486	0.168351	0.2445935	0.219467
## 3301	-1.423e-01	0.2397182	0.148849	0.412450	0.1624531	0.356021
## 3302	7.792e-02	0.1216489	0.446172	0.237364	0.4025120	0.111751
## 3303	-8.344e-02	-0.0124946	0.143989	0.155074	0.4419442	0.335382
## 3304	-4.643e-03	0.0353021	0.094194	0.205382	0.1839314	0.324818
## 3305	1.432e-01	-0.0155643	0.157638	0.153217	0.2381919	0.182169
## 3306	1.047e-01	-0.0479580	0.099565	0.261350	0.3046537	0.225255
## 3307	3.693e-02	-0.0581603	0.210578	0.442799	0.5869108	0.286523
## 3308	7.179e-02	-0.0352735	0.353678	0.438042	0.5665383	0.298749

## 3309	1.665e-01	-0.2734380	0.222824	0.411442	0.8050133	-0.062177
## 3310	2.176e-03	0.0194135	0.278986	0.383224	0.2486587	-0.114540
## 3311	-3.929e-01	-0.2042468	0.029676	0.410712	0.4485353	0.349828
## 3312	5.849e-02	0.0002726	0.131492	0.362417	0.3674911	0.155455
## 3313	2.960e-01	-0.0339303	0.128463	-0.150844	0.5500087	0.276019
## 3314	4.348e-02	0.0767198	0.087098	0.291240	0.2834899	-0.412141
## 3315	3.107e-01	-0.0381125	0.178013	0.179099	0.4487575	0.120185
## 3316	8.744e-02	-0.0993189	0.342538	-0.121741	0.3154417	0.645999
## 3317	7.463e-01	-0.2410918	0.333762	0.713539	-0.6205338	-0.917391
## 3318	-2.874e-02	0.1269149	0.305349	0.303857	0.3834475	0.134907
## 3319	-6.583e-02	-0.0863497	0.340394	0.257087	0.9219045	-0.228222
## 3320	-1.688e-01	0.1791013	0.047312	0.706427	0.0079641	-0.532912
## 3321	-1.804e-02	-0.0555066	0.099994	0.171753	0.0966854	0.432978
## 3322	1.485e-02	-0.1841752	0.236963	0.289537	0.7818788	0.810339
## 3400	6.458e-02	-0.0777015	0.231940	0.332273	0.1784571	0.275088
## 3401	NA	NA	NA	NA	NA	NA
## 3402	NA	NA	NA	NA	NA	NA
## 3403	NA	NA	NA	NA	NA	NA
## 3404	NA	NA	NA	NA	NA	NA
## 3500	-7.712e-03	-0.0870301	0.114752	0.705367	0.5551960	0.595854
## 3501	NA	NA	NA	NA	NA	NA
## 3502	NA	NA	NA	NA	NA	NA
## 3503	NA	NA	NA	NA	NA	NA
## 3504	NA	NA	NA	NA	NA	NA
## 3505	NA	NA	NA	NA	NA	NA
## 3506	NA	NA	NA	NA	NA	NA
## 3600	NA	NA	NA	NA	NA	NA
## 3601	NA	NA	NA	NA	NA	NA
## 3602	NA	NA	NA	NA	NA	NA
## 3603	NA	NA	NA	NA	NA	NA
## 3604	NA	NA	NA	NA	NA	NA
## 3605	NA	NA	NA	NA	NA	NA
## 3606	NA	NA	NA	NA	NA	NA
## 3607	NA	NA	NA	NA	NA	NA
## 3608	NA	NA	NA	NA	NA	NA
## 3609	NA	NA	NA	NA	NA	NA
## 3610	NA	NA	NA	NA	NA	NA
## 3611	-4.818e-01	0.1719048	0.477462	0.395975	1.0219797	0.549907
## 3612	-1.391e-01	-0.0629172	0.248944	0.291578	0.2530231	0.437308
## 3613	NA	NA	NA	NA	NA	NA
## 3614	NA	NA	NA	NA	NA	NA
## 3615	NA	NA	NA	NA	NA	NA
## 3616	-1.841e-01	0.2014220	-0.013167	-0.033394	0.4254274	-0.224416
##	FFA2	FLA2	FFA3	FLA4		
## 1000	-4.761e-02	2.296e-02	-0.0422134	0.0081127		
## 1100	8.535e-02	-3.153e-02	0.0768362	0.0014730		
## 1101	2.743e-01	4.079e-02	0.2796293	0.0892749		
## 1102	1.004e-02	-1.679e-02	0.0083640	-0.0156223		
## 1103	4.966e-03	1.665e-02	0.0085176	0.0180803		
## 1104	-1.124e-02	4.898e-03	-0.0102790	0.0023916		

## 1105	1.593e-03	-2.366e-02	-0.0023729	-0.0233333
## 1106	1.353e-01	-3.762e-02	0.1323229	-0.0252902
## 1107	-6.652e-02	-7.168e-02	-0.0908285	-0.0898448
## 1108	-9.582e-02	-1.167e-02	-0.0969830	-0.0214331
## 1109	5.106e-02	-1.200e-01	0.0111478	-0.1029600
## 1110	1.027e-02	1.791e-02	0.0126178	0.0194082
## 1111	-7.282e-02	7.729e-03	-0.0721891	-0.0011049
## 1200	NA	NA	NA	NA
## 1201	1.407e-02	4.010e-02	0.0396988	0.0488292
## 1202	-1.421e-01	1.924e-01	0.0329185	0.0633861
## 1203	6.851e-02	2.587e-02	0.0903227	0.0831432
## 1204	1.542e-01	-5.963e-02	0.1215606	0.0353329
## 1205	NA	NA	NA	NA
## 1206	NA	NA	NA	NA
## 1207	3.143e-01	3.088e-02	0.3361964	0.2506186
## 1208	7.704e-01	-8.237e-01	-0.0150959	-0.0878509
## 1209	NA	NA	NA	NA
## 1210	9.266e-01	-3.939e-01	0.6056491	0.5015622
## 1211	-2.969e-01	3.693e-01	0.0588340	0.0839755
## 1212	-2.661e-01	2.538e-01	-0.0344757	0.0031621
## 1213	-4.390e-03	6.404e-02	0.0555306	0.0616383
## 1300	1.472e-01	-1.595e-01	0.1008634	-0.1054570
## 1301	NA	NA	NA	NA
## 1302	NA	NA	NA	NA
## 1303	3.901e-02	-1.335e-02	0.0381136	-0.0097975
## 1304	8.193e-02	-1.297e-01	0.0681074	-0.1167167
## 1305	-1.231e-04	-7.239e-02	-0.0060061	-0.0726135
## 1306	-2.540e-02	-7.079e-02	-0.0336503	-0.0740418
## 1307	5.320e-02	1.193e-02	0.0531098	0.0111923
## 1308	-7.582e-02	-5.727e-02	-0.0878273	-0.0803117
## 1309	4.444e-02	-6.207e-02	0.0349764	-0.0572511
## 1310	-2.805e-02	3.349e-02	-0.0269572	0.0320803
## 1311	2.278e-02	1.442e-02	0.0239635	0.0169089
## 1312	2.063e-02	-3.714e-02	0.0186254	-0.0356888
## 1313	-1.008e-02	-1.210e-01	-0.0209047	-0.1221622
## 1314	-4.910e-02	-4.173e-02	-0.0523773	-0.0472621
## 1315	1.534e-01	-2.898e-01	0.1182335	-0.2636285
## 1400	3.261e-02	3.849e-01	0.2320861	0.3997156
## 1401	-1.637e-01	1.278e-01	-0.1048589	0.0371354
## 1402	-3.014e-02	6.784e-02	-0.0030378	0.0549575
## 1403	-2.192e-02	4.566e-02	0.0067794	0.0317165
## 1404	-3.798e-01	1.244e-01	-0.2913860	-0.0545038
## 1405	-3.796e-02	1.181e-01	0.0346664	0.0911763
## 1406	-1.492e-01	1.693e-01	-0.0727411	0.1031192
## 1407	8.875e-03	1.763e-01	0.1242402	0.1822349
## 1408	-2.341e-02	3.155e-02	-0.0057506	0.0178141
## 1409	1.375e-01	3.307e-02	0.1546468	0.1039314
## 1410	-1.361e-02	1.080e-01	0.0650681	0.0902636
## 1500	1.101e-02	-2.159e-01	-0.0326164	-0.2140460
## 1501	3.373e-01	-2.182e-02	0.3370574	-0.0112988

## 1502	8.259e-02	-3.845e-01	0.0574853	-0.3743702
## 1503	-1.100e-01	-9.659e-02	-0.1198321	-0.1073715
## 1504	NA	NA	NA	NA
## 1505	2.860e-01	2.300e-02	0.2832173	0.0167792
## 1506	NA	NA	NA	NA
## 1507	NA	NA	NA	NA
## 1508	NA	NA	NA	NA
## 1600	7.041e-02	2.581e-02	0.0727426	0.0334212
## 1601	NA	NA	NA	NA
## 1602	3.319e-02	-2.035e-02	0.0319682	-0.0174754
## 1603	1.451e-01	-1.001e-01	0.1122831	-0.0416032
## 1604	-3.925e-03	2.140e-02	-0.0056580	0.0217369
## 1605	3.471e-02	-2.053e-02	0.0344664	-0.0200496
## 1606	8.471e-02	3.756e-02	0.0849883	0.0384388
## 1607	1.169e-01	2.851e-01	0.1694644	0.3169618
## 1700	3.374e-01	-1.800e-01	0.2686083	-0.0114277
## 1701	NA	NA	NA	NA
## 1702	2.364e-01	9.643e-02	0.3172622	0.2712898
## 1703	1.108e-03	2.835e-01	0.0895489	0.2832832
## 1704	NA	NA	NA	NA
## 1705	-6.429e-02	1.286e-01	0.0241667	0.0881405
## 1706	1.325e-01	-1.417e-01	0.0833793	-0.0966279
## 1707	2.844e-01	-1.421e+00	-0.0781654	-1.1370000
## 1708	2.154e-01	-2.135e-01	0.1664241	-0.1369851
## 1709	2.516e-01	-4.910e-01	0.0085983	-0.3026808
## 1710	1.041e-01	2.425e-02	0.1139099	0.0721179
## 1711	7.565e-01	-5.002e-01	NA	-0.3781017
## 1712	3.327e-01	-1.857e-01	0.3095388	-0.1205379
## 1800	NA	NA	NA	NA
## 1801	NA	NA	NA	NA
## 1802	-1.975e-02	8.478e-02	0.0204504	0.0733700
## 1803	5.799e-02	1.406e-01	0.1278710	0.1687126
## 1804	7.859e-02	4.064e-03	0.0795988	0.0232813
## 1900	1.700e-01	-1.216e-01	0.0974409	-0.0178127
## 1901	-4.018e-02	-1.418e-02	-0.0427542	-0.0234564
## 1902	1.481e-02	-1.035e-01	-0.0280752	-0.0954917
## 1903	NA	NA	NA	NA
## 1904	-1.731e-01	1.448e-01	-0.0923949	0.0570297
## 1905	NA	NA	NA	NA
## 1906	9.678e-02	-3.726e-02	0.0915840	-0.0200772
## 1907	-4.489e-02	-2.052e-03	-0.0454181	-0.0148016
## 1908	1.618e-02	-1.649e-02	0.0136147	-0.0132275
## 1909	1.304e-01	1.342e-01	0.1831836	0.1734026
## 1910	5.182e-02	2.954e-02	NA	0.0482513
## 1911	6.764e-02	-4.667e-02	0.0422570	-0.0154064
## 1912	2.543e-01	-1.000e-01	0.2503986	-0.0881718
## 1913	5.061e-02	-7.801e-02	0.0477750	-0.0737078
## 2000	2.376e-01	3.768e-01	0.4588514	0.5084119
## 2001	-7.543e-01	5.170e-01	-0.3686489	-0.0620629
## 2002	3.354e-02	5.259e-02	0.0649798	0.0719899

##	2003	4.721e-02	-3.565e-02	0.0323232	-0.0129477
##	2100	7.661e-02	-5.754e-03	0.0763333	0.0031958
##	2101	NA	NA	NA	NA
##	2102	6.893e-01	-2.927e-01	0.6310451	-0.1855100
##	2103	-8.593e-03	-2.836e-01	-0.0635853	-0.2893353
##	2104	NA	NA	NA	NA
##	2105	1.385e-01	-4.270e-01	-0.0696453	-0.3426277
##	2200	2.094e-01	-1.783e-01	0.1301472	-0.0518218
##	2201	-5.987e-02	-9.615e-02	-0.1329411	-0.1560708
##	2202	-1.454e-01	-1.665e-01	-0.2502876	-0.2292129
##	2203	NA	NA	NA	NA
##	2204	6.592e-02	-3.594e-01	-0.0221568	-0.3252804
##	2205	1.783e-03	2.084e-03	0.0026885	0.0027617
##	2206	NA	NA	NA	NA
##	2207	7.773e-01	-4.190e-01	0.6899772	-0.2041479
##	2208	1.678e-01	-4.490e-01	0.0235817	-0.3801867
##	2209	2.850e-01	-2.173e-01	0.1956173	-0.1207527
##	2210	2.621e-01	-4.454e-01	0.1157941	-0.3385182
##	2211	2.152e-01	-3.673e-01	0.0892451	-0.2474849
##	2212	7.901e-01	-3.261e-01	0.6401019	0.0216543
##	2213	-1.156e-01	-1.293e-01	-0.1582425	-0.1781420
##	2214	2.933e-01	1.421e-01	0.3252940	0.2184830
##	2215	2.720e-01	-3.382e-01	-0.0481269	-0.1727227
##	2216	NA	NA	NA	NA
##	2300	1.612e-01	-1.224e-01	0.1083161	-0.0462167
##	2301	-2.191e-02	1.215e-01	0.0449835	0.1096615
##	2302	-7.471e-02	1.784e-01	-0.0085176	0.1573365
##	2303	2.030e-02	-4.837e-02	0.0132923	-0.0448433
##	2304	1.105e-02	2.233e-02	0.0120793	0.0228614
##	2305	-7.137e-02	8.268e-02	-0.0553723	0.0555168
##	2306	2.449e-02	-3.665e-03	0.0242270	0.0017079
##	2307	-5.656e-02	1.345e-01	-0.0314724	0.1211995
##	2308	9.324e-02	-7.117e-02	0.0626020	-0.0264765
##	2309	1.611e-02	-8.255e-02	-0.0011331	-0.0780920
##	2310	8.781e-02	1.162e-01	0.0925939	0.1181923
##	2311	-7.088e-03	2.608e-01	0.0616611	0.2576118
##	2312	-1.048e-01	5.659e-02	-0.0900096	0.0292904
##	2400	1.935e-01	8.257e-02	0.2143153	0.1064753
##	2401	NA	NA	NA	NA
##	2402	-6.190e-03	-3.974e-02	-0.0029797	-0.0393382
##	2403	6.556e-02	2.319e-02	0.0676704	0.0313981
##	2404	6.591e-03	7.461e-02	0.0107693	0.0750881
##	2405	NA	NA	0.0342393	NA
##	2406	-9.887e-02	7.888e-02	-0.1056149	0.0912263
##	2500	1.437e-01	6.004e-02	0.1514929	0.0839161
##	2501	NA	NA	NA	NA
##	2502	2.167e-01	-2.494e-01	0.1953140	-0.2175862
##	2503	3.237e-02	-1.747e-01	-0.0141581	-0.1649859
##	2504	-1.283e-02	3.573e-02	-0.0046148	0.0337564
##	2505	5.060e-02	-4.789e-02	0.0516691	-0.0496402

## 2506	1.356e-01	-5.726e-02	0.1257076	-0.0311086
## 2507	7.810e-02	-2.741e-02	0.0786406	-0.0176295
## 2508	8.595e-02	4.985e-02	0.0918474	0.0584667
## 2600	2.020e-01	1.976e-01	0.3173530	0.3119370
## 2601	2.684e-01	-2.174e-02	0.2596031	0.1482102
## 2602	1.323e-01	3.742e-01	0.2583757	0.5065000
## 2603	NA	NA	NA	NA
## 2604	1.896e-01	-1.580e-01	0.1542511	-0.1026889
## 2605	2.524e-01	7.455e-02	0.2658715	0.1330622
## 2606	NA	NA	NA	NA
## 2607	1.168e-01	-3.817e-01	0.0197761	-0.3402009
## 2608	NA	NA	0.2418841	0.2421100
## 2609	-2.808e-01	8.011e-01	0.3364365	0.5972339
## 2610	-1.293e+00	5.125e-01	-0.7804243	-0.1870023
## 2611	3.954e-01	-2.814e-01	0.3295092	-0.1566309
## 2612	8.397e-01	-3.993e-01	0.7944557	-0.3094766
## 2613	2.451e-01	1.778e-01	0.2842791	0.2389161
## 2614	1.366e-01	-2.735e-01	0.0275739	-0.2136740
## 2700	1.197e-01	-1.180e-01	0.0872808	-0.0799063
## 2701	7.105e-02	5.211e-02	0.0875014	0.0736577
## 2702	7.108e-02	1.427e-01	0.1020851	0.1612694
## 2703	1.260e-01	1.424e-01	0.1731431	0.1953643
## 2704	8.535e-03	-8.061e-02	-0.0081129	-0.0796754
## 2705	4.596e-02	1.055e-01	0.0554176	0.1106319
## 2706	2.468e-01	6.354e-02	0.2487860	0.0892436
## 2707	-1.009e-02	-1.151e-02	-0.0114129	-0.0116610
## 2708	1.335e-01	1.054e-01	0.1573025	0.1561984
## 2709	NA	NA	NA	NA
## 2710	-2.743e-02	-1.163e-01	-0.0401368	-0.1222270
## 2711	-1.011e-01	-3.594e-02	-0.1047292	-0.0514214
## 2712	-1.166e-01	9.783e-02	-0.0987569	0.0701023
## 2713	5.527e-02	4.662e-04	0.0555561	0.0005161
## 2714	3.659e-01	-7.035e-03	0.3932183	0.1872616
## 2715	1.215e-01	-1.989e-01	0.0893607	-0.1753337
## 2716	1.247e-01	1.727e-01	0.1622212	0.2021833
## 2717	-3.312e-02	2.626e-02	-0.0274615	0.0200460
## 2718	-5.630e-02	2.406e-01	-0.0172958	0.2306352
## 2719	1.162e-01	-2.378e-02	0.1071704	0.0202320
## 2720	9.638e-02	-6.316e-02	0.0919482	-0.0592491
## 2721	1.805e-01	3.623e-03	0.1812342	-0.0006603
## 2722	1.388e-01	-3.884e-02	0.1330304	-0.0336827
## 2723	5.839e-02	-7.100e-02	0.0487858	-0.0606426
## 2724	-2.819e-02	1.558e-01	-0.0184571	0.1534389
## 2725	-1.330e-02	3.570e-02	-0.0124921	0.0352082
## 2726	9.429e-03	-1.158e-01	0.0013808	-0.1155762
## 2727	2.655e-01	-1.455e-03	0.2655933	0.0305290
## 2728	8.600e-02	4.911e-02	0.0916488	0.0599367
## 2729	1.474e-01	-3.770e-02	0.1402382	-0.0106115
## 2730	4.848e-02	4.652e-02	0.0532962	0.0524428
## 2731	9.179e-02	-7.156e-02	0.0944112	-0.0766991

##	2732	-4.282e-02	-3.859e-02	-0.0569579	-0.0551429
##	2733	-8.085e-02	1.295e-01	-0.0527962	0.1100465
##	2734	1.700e-01	-9.024e-02	0.1305990	-0.0138949
##	2735	-4.690e-02	5.694e-02	-0.0405254	0.0514906
##	2736	4.671e-02	-1.086e-01	0.0002273	-0.0899695
##	2737	-2.523e-02	8.627e-03	-0.0232588	0.0031886
##	2738	1.023e-01	2.627e-02	0.1072647	0.0454942
##	2739	-1.758e-02	-3.651e-02	-0.0274503	-0.0412511
##	2740	1.254e-01	1.050e-01	0.1212963	0.0966086
##	2741	1.615e-02	2.274e-02	0.0211200	0.0273254
##	2742	-2.489e-01	-1.650e-01	-0.3069058	-0.2463537
##	2743	1.129e-01	1.566e-02	0.1121395	0.0106037
##	2744	NA	NA	NA	NA
##	2745	5.622e-01	4.914e-01	0.4138998	0.2013064
##	2746	-7.749e-02	-1.223e-02	-0.0776302	-0.0139709
##	2747	4.827e-02	3.068e-01	0.0614514	0.3107755
##	2748	1.058e-01	4.938e-02	0.1163017	0.0746775
##	2800	7.790e-02	-6.421e-02	0.0725402	-0.0547539
##	2801	NA	NA	NA	NA
##	2802	1.025e-02	2.262e-03	0.0106153	0.0031490
##	2803	4.004e-02	6.432e-02	0.0524779	0.0748145
##	2804	9.181e-02	7.715e-02	0.0885474	0.0629531
##	2805	3.434e-02	-2.266e-01	-0.0397069	-0.2152365
##	2806	1.186e-02	7.743e-05	0.0119068	-0.0004429
##	2807	NA	NA	NA	NA
##	2808	2.025e-01	5.959e-02	0.2058496	0.0669527
##	2809	-1.109e-01	-6.840e-02	-0.1176622	-0.0862321
##	2900	-2.311e-01	-2.039e-01	-0.3583238	-0.3293164
##	2901	NA	NA	NA	NA
##	2902	8.941e-05	-7.334e-02	-0.0339634	-0.0797026
##	2903	NA	NA	NA	NA
##	2904	NA	NA	NA	NA
##	2905	5.894e-02	4.778e-01	0.0985294	0.5135579
##	2906	NA	NA	NA	NA
##	2907	NA	NA	NA	NA
##	2908	NA	NA	NA	NA
##	2909	NA	NA	NA	NA
##	2910	3.254e-01	-1.682e-01	0.1880689	0.1032923
##	2911	NA	NA	NA	NA
##	2912	NA	NA	NA	NA
##	2913	NA	NA	NA	NA
##	2914	NA	NA	NA	NA
##	2915	NA	NA	NA	NA
##	2916	1.695e-02	-1.470e-02	0.0137673	-0.0112928
##	2917	NA	NA	NA	NA
##	2918	NA	NA	NA	NA
##	2919	NA	NA	NA	NA
##	2920	NA	NA	NA	NA
##	2921	4.326e-01	1.702e-01	0.4747527	0.2873589
##	2922	NA	NA	NA	NA

## 2923	NA	NA	NA	NA
## 3000	-3.933e-01	-3.915e-01	-0.6656160	-0.5856122
## 3001	NA	NA	NA	NA
## 3002	6.877e-02	1.368e-02	0.0700558	0.0226872
## 3003	-4.284e-02	-3.186e-02	-0.0503110	-0.0425071
## 3004	8.418e-02	-1.181e-01	0.0521763	-0.0905928
## 3005	8.318e-02	1.915e-01	0.1207462	0.2089830
## 3100	7.541e-02	5.466e-02	0.0771317	0.0608410
## 3101	-5.554e-01	-2.689e-01	-0.4812021	-0.0831995
## 3102	NA	NA	0.0414729	0.0414729
## 3103	-3.118e-01	5.673e-02	-0.2774217	-0.0985828
## 3104	2.008e-01	-3.453e-02	0.1937273	0.0208020
## 3105	3.960e-01	-5.532e-01	0.0569094	-0.3409294
## 3106	6.359e-02	4.881e-02	0.0642526	0.0431816
## 3107	-2.082e-01	5.166e-02	-0.1957947	-0.0168630
## 3108	NA	NA	NA	NA
## 3109	-7.394e-01	7.237e-01	NA	-0.0158251
## 3110	1.504e-01	4.269e-01	0.1459418	0.4263565
## 3200	2.811e-02	-4.232e-02	NA	-0.0342337
## 3201	4.293e-01	-1.666e-01	0.3422606	0.0284354
## 3202	-5.641e-03	6.129e-02	0.0166625	0.0605979
## 3203	-5.359e-02	-1.269e-01	-0.0705491	-0.1343282
## 3204	-3.509e-02	-1.731e-02	-0.0428428	-0.0321389
## 3205	-7.172e-02	6.021e-02	-0.0532741	0.0349000
## 3206	1.877e-01	-2.967e-02	0.1787056	0.0419491
## 3207	-6.208e-02	-5.091e-02	-0.0899136	-0.0843967
## 3300	1.271e-02	-2.055e-02	-0.0020978	-0.0117746
## 3301	-1.172e-01	2.193e-01	0.0297531	0.1417151
## 3302	1.939e-01	-9.180e-02	0.1353031	0.0519374
## 3303	-7.343e-02	-1.098e-02	-0.0803232	-0.0574478
## 3304	3.025e-03	2.864e-02	0.0220686	0.0306679
## 3305	1.546e-01	-2.700e-02	0.1364524	0.0785392
## 3306	1.139e-01	-1.106e-01	0.0676071	-0.0635238
## 3307	4.254e-02	-5.715e-02	0.0316274	-0.0400947
## 3308	1.292e-01	-3.053e-02	0.1095654	0.0549280
## 3309	1.110e-01	-2.022e-01	-0.0172433	-0.1362284
## 3310	2.960e-02	3.291e-02	0.0563546	0.0571148
## 3311	-3.410e-01	-1.508e-01	-0.4106678	-0.3137493
## 3312	1.172e-01	-3.812e-02	0.0897984	0.0489852
## 3313	2.757e-01	6.264e-02	0.2975771	0.1518632
## 3314	6.538e-02	5.846e-02	0.1117013	0.1117689
## 3315	2.724e-01	2.987e-02	0.2949988	0.2337590
## 3316	6.988e-02	-6.398e-02	0.0139515	-0.0025299
## 3317	8.589e-01	-3.305e-01	0.5656003	0.3660254
## 3318	-4.180e-02	5.209e-02	-0.0061282	0.0263915
## 3319	-7.728e-02	-1.120e-01	-0.1318530	-0.1569633
## 3320	-1.357e-01	1.745e-01	0.0163742	0.0660892
## 3321	3.903e-02	-1.020e-01	-0.0298530	-0.0754451
## 3322	4.496e-02	-1.855e-01	-0.0974246	-0.1528699
## 3400	7.548e-02	-6.825e-02	0.0796347	-0.0719356

```

## 3401      NA      NA      NA      NA
## 3402      NA      NA      NA      NA
## 3403      NA      NA      NA      NA
## 3404      NA      NA      NA      NA
## 3500  6.605e-02  3.346e-02  0.0837957  0.0677439
## 3501      NA      NA      NA      NA
## 3502      NA      NA      NA      NA
## 3503      NA      NA      NA      NA
## 3504 -3.155e-01  2.408e-01 -0.1952973  0.1526518
## 3505      NA      NA      NA      NA
## 3506      NA      NA      NA      NA
## 3600  2.638e-01 -2.491e-01  0.1817295 -0.1640844
## 3601      NA      NA      NA      NA
## 3602      NA      NA      NA      NA
## 3603      NA      NA      NA      NA
## 3604      NA      NA      NA      NA
## 3605      NA      NA      NA      NA
## 3606      NA      NA      NA      NA
## 3607  0.000e+00  0.000e+00  0.0000000  0.0000000
## 3608      NA      NA      NA      NA
## 3609      NA      NA      NA      NA
## 3610      NA      NA      NA      NA
## 3611 -2.951e-01  1.314e-01 -0.2822313  0.1020236
## 3612 -1.231e-01 -1.072e-01 -0.1632496 -0.1605523
## 3613      NA      NA      NA      NA
## 3614      NA      NA      NA      NA
## 3615      NA      NA      NA      NA
## 3616 -1.963e-01  1.672e-01 -0.1244881  0.0856803

```

```
print(RegP)
```

```

##      FFA1p      FLA1p      2p      3p      4p      5+p
## 1000 1.121e-01 0.8726067 8.211e-04 7.987e-05 1.718e-03 1.396e-06
## 1100 2.643e-01 0.5419315 7.668e-03 4.699e-02 2.445e-03 6.508e-04
## 1101 1.042e-01 0.9374396 1.639e-01 1.942e-01 7.210e-03 2.770e-02
## 1102 9.196e-01 0.9586209 6.112e-01 4.902e-02 2.038e-03 2.890e-03
## 1103 1.602e-01 0.3359658 9.499e-08 2.522e-08 9.329e-09 1.072e-11
## 1104 3.914e-01 0.9288679 3.919e-04 2.670e-03 3.275e-03 1.136e-06
## 1105 4.475e-01 0.2026996 1.283e-10 1.491e-09 1.081e-11 2.290e-13
## 1106 6.842e-03 0.3220507 8.441e-05 8.189e-07 7.247e-05 3.406e-08
## 1107 3.341e-01 0.1349871 1.524e-02 2.994e-02 7.841e-04 6.143e-05
## 1108 5.952e-02 0.9715112 6.108e-01 6.205e-02 3.738e-02 3.868e-04
## 1109 3.874e-01 0.0481734 2.957e-02 9.541e-03 5.712e-03 6.242e-03
## 1110 7.593e-01 0.6951165 1.502e-05 2.019e-06 5.355e-13 3.108e-17
## 1111 8.530e-02 0.8779153 1.144e-01 4.334e-01 7.911e-02 3.418e-02
## 1200      NA      NA      NA      NA      NA      NA
## 1201 8.439e-01 0.4163319 8.527e-02 2.528e-01 4.180e-03 3.542e-03
## 1202 4.241e-01 0.3043018 1.582e-01 4.856e-02 8.401e-01 3.378e-06
## 1203 7.923e-01 0.8560898 9.710e-04 5.634e-03 5.591e-02 4.438e-01
## 1204 4.504e-01 0.5119481 9.529e-05 1.956e-01 6.887e-01 2.697e-01

```

## 1205	NA	NA	NA	NA	NA	NA
## 1206	NA	NA	NA	NA	NA	NA
## 1207	1.546e-01	0.2324053	1.748e-02	1.851e-04	5.537e-06	3.687e-04
## 1208	3.730e-01	0.3328062	2.088e-01	8.784e-10	3.796e-01	5.741e-01
## 1209	NA	NA	NA	NA	NA	NA
## 1210	NA	NA	NA	NA	NA	NA
## 1211	3.172e-01	0.2002013	6.214e-03	1.837e-02	7.166e-27	4.242e-01
## 1212	9.195e-02	0.0895296	2.756e-01	2.962e-08	1.622e-08	8.546e-01
## 1213	4.759e-01	0.5374429	2.906e-01	2.315e-01	2.184e-01	5.659e-04
## 1300	6.013e-02	0.0307581	2.288e-01	9.192e-01	3.043e-02	9.311e-03
## 1301	NA	NA	NA	NA	NA	NA
## 1302	NA	NA	NA	NA	NA	NA
## 1303	3.139e-01	0.6393854	2.028e-02	5.971e-03	5.816e-04	1.604e-07
## 1304	1.431e-01	0.1347738	4.050e-01	5.245e-01	3.081e-01	6.162e-01
## 1305	7.694e-01	0.2463781	4.916e-01	2.665e-01	2.229e-02	3.589e-02
## 1306	6.108e-01	0.4708135	1.718e-01	9.413e-01	1.979e-01	9.414e-01
## 1307	2.121e-01	0.5151401	3.229e-01	1.949e-01	4.756e-02	1.246e-02
## 1308	1.039e-01	0.2827678	3.460e-06	5.445e-08	8.426e-10	8.574e-14
## 1309	6.476e-01	0.3626316	7.401e-01	3.760e-01	9.468e-01	2.847e-01
## 1310	5.265e-01	0.4796031	2.131e-01	1.045e-01	5.609e-02	1.784e-02
## 1311	9.984e-01	0.8461870	2.956e-02	4.568e-03	5.589e-03	2.742e-04
## 1312	9.790e-01	0.1327837	2.643e-03	1.443e-03	6.940e-04	8.952e-05
## 1313	6.563e-01	0.3222659	2.350e-01	1.485e-01	4.073e-02	1.044e-02
## 1314	1.894e-01	0.0572572	3.146e-01	1.636e-01	3.324e-03	5.987e-04
## 1315	1.243e-01	0.1765943	2.158e-01	2.099e-01	1.989e-01	6.450e-01
## 1400	8.698e-01	0.0040333	3.876e-03	7.872e-03	2.712e-04	6.477e-01
## 1401	2.256e-01	0.6362315	3.386e-01	3.022e-01	8.354e-01	7.041e-02
## 1402	6.308e-01	0.2855946	4.978e-03	1.886e-02	5.302e-13	3.835e-03
## 1403	4.112e-01	0.6539586	2.585e-04	4.198e-01	5.525e-03	8.875e-01
## 1404	NA	NA	NA	NA	NA	NA
## 1405	4.739e-01	0.2567773	1.053e-01	8.400e-02	7.307e-01	7.602e-01
## 1406	4.060e-02	0.0384575	4.243e-02	8.575e-01	4.967e-04	6.535e-01
## 1407	9.425e-01	0.1225286	2.302e-01	6.197e-02	2.302e-01	1.930e-01
## 1408	3.525e-01	0.5343427	6.431e-04	5.240e-02	1.912e-01	2.650e-14
## 1409	7.386e-02	0.7663806	2.939e-03	5.805e-01	3.026e-01	8.515e-01
## 1410	9.618e-01	0.8665850	5.924e-03	2.706e-01	9.014e-01	1.840e-02
## 1500	7.821e-01	0.0002910	2.686e-03	6.500e-04	1.132e-04	9.521e-05
## 1501	NA	NA	NA	NA	NA	NA
## 1502	1.697e-01	0.0008660	5.134e-01	6.444e-01	9.223e-01	8.910e-01
## 1503	2.005e-01	0.2335271	6.892e-01	9.780e-01	8.780e-01	7.103e-01
## 1504	NA	NA	NA	NA	NA	NA
## 1505	NA	NA	NA	NA	NA	NA
## 1506	NA	NA	NA	NA	NA	NA
## 1507	NA	NA	NA	NA	NA	NA
## 1508	NA	NA	NA	NA	NA	NA
## 1600	2.757e-01	0.6842562	6.026e-01	5.228e-01	5.342e-01	2.082e-02
## 1601	NA	NA	NA	NA	NA	NA
## 1602	6.458e-01	0.9099878	2.333e-01	4.938e-02	2.403e-01	1.755e-02
## 1603	1.251e-01	0.1242391	4.485e-02	2.078e-02	4.385e-01	1.255e-03
## 1604	6.006e-01	0.4740872	7.434e-01	4.741e-01	5.801e-01	8.608e-01

## 1605	2.105e-01	0.6584344	6.238e-01	8.719e-01	6.330e-01	2.694e-01
## 1606	8.370e-02	0.4750786	6.961e-02	2.553e-02	1.660e-01	4.617e-02
## 1607	2.368e-01	0.0055489	4.036e-02	6.993e-02	2.258e-02	6.816e-02
## 1700	4.344e-02	0.4151009	3.313e-01	1.631e-01	5.359e-01	3.428e-01
## 1701	NA	NA	NA	NA	NA	NA
## 1702	4.008e-01	0.6681695	9.850e-02	9.796e-02	5.235e-01	7.173e-01
## 1703	1.245e-01	0.1501092	5.529e-02	1.324e-01	2.478e-04	7.038e-01
## 1704	NA	NA	NA	NA	NA	NA
## 1705	9.897e-01	0.7614769	8.714e-01	7.070e-01	7.320e-01	1.091e-01
## 1706	1.428e-01	0.1063904	5.083e-01	3.786e-01	8.656e-01	7.906e-01
## 1707	NA	NA	NA	NA	NA	NA
## 1708	NA	NA	NA	NA	NA	NA
## 1709	2.409e-01	0.2558847	3.576e-01	4.023e-01	6.250e-02	7.224e-01
## 1710	2.961e-01	0.7116339	1.705e-01	6.282e-01	7.544e-01	4.959e-01
## 1711	NA	NA	NA	NA	NA	NA
## 1712	6.510e-02	0.0628202	1.421e-02	2.926e-02	5.425e-03	2.854e-03
## 1800	NA	NA	NA	NA	NA	NA
## 1801	NA	NA	NA	NA	NA	NA
## 1802	4.424e-01	0.3926748	2.668e-01	3.497e-02	9.330e-01	5.291e-01
## 1803	9.121e-01	0.3242163	3.056e-01	6.166e-01	5.781e-01	7.205e-03
## 1804	8.478e-01	0.2792841	3.444e-02	4.408e-07	8.306e-02	9.720e-01
## 1900	1.019e-01	0.3550160	4.086e-04	1.567e-05	2.213e-02	1.703e-03
## 1901	4.954e-01	0.9337366	4.152e-01	5.552e-01	1.311e-01	4.275e-01
## 1902	9.339e-01	0.5452833	3.977e-01	3.538e-01	7.571e-01	8.873e-03
## 1903	NA	NA	NA	NA	NA	NA
## 1904	9.268e-03	0.0132455	8.313e-01	6.158e-01	3.429e-01	9.586e-01
## 1905	NA	NA	NA	NA	NA	NA
## 1906	6.594e-02	0.3563101	1.087e-01	5.933e-01	9.449e-01	2.340e-01
## 1907	3.806e-01	0.9736011	2.297e-01	2.612e-01	2.641e-01	1.405e-01
## 1908	7.019e-01	0.7418328	5.389e-01	3.021e-01	6.499e-01	2.782e-02
## 1909	4.658e-01	0.1393235	7.078e-01	6.828e-01	8.173e-02	3.296e-03
## 1910	2.428e-01	0.5729675	6.359e-02	3.330e-01	1.047e-02	6.716e-03
## 1911	2.880e-01	0.3468299	2.227e-01	9.045e-01	6.452e-01	8.133e-03
## 1912	1.070e-01	0.5980736	2.037e-03	1.079e-01	1.845e-02	1.045e-01
## 1913	NA	NA	NA	NA	NA	NA
## 2000	1.976e-01	0.0377051	3.135e-01	5.633e-02	3.676e-07	7.478e-01
## 2001	5.740e-06	0.0003954	1.409e-02	5.349e-01	1.907e-04	3.804e-02
## 2002	7.820e-01	0.7294008	1.721e-05	5.354e-04	1.218e-03	6.515e-10
## 2003	9.089e-01	0.8894618	1.979e-04	1.128e-02	6.759e-01	2.763e-02
## 2100	6.436e-01	0.8913663	6.394e-02	3.394e-01	1.550e-01	3.151e-03
## 2101	NA	NA	NA	NA	NA	NA
## 2102	4.508e-03	0.6197057	2.050e-03	9.909e-01	7.376e-01	1.140e-01
## 2103	NA	NA	NA	NA	NA	NA
## 2104	NA	NA	NA	NA	NA	NA
## 2105	3.555e-01	0.0198010	6.936e-01	1.333e-01	9.655e-01	3.244e-01
## 2200	1.914e-01	0.3881582	7.296e-01	3.787e-01	8.728e-01	4.823e-01
## 2201	9.407e-01	0.9648303	4.680e-01	3.713e-01	8.181e-01	7.445e-02
## 2202	8.961e-01	0.3826373	2.859e-01	2.822e-01	4.347e-01	9.178e-01
## 2203	NA	NA	NA	NA	NA	NA
## 2204	5.751e-01	0.1065851	4.797e-01	3.553e-01	3.706e-01	4.209e-01

##	2205	9.502e-01	0.9424609	3.337e-02	8.062e-01	7.410e-05	3.036e-02
##	2206	NA	NA	NA	NA	NA	NA
##	2207	1.480e-03	0.1234897	2.701e-01	4.155e-01	1.508e-06	5.909e-01
##	2208	1.959e-01	0.0007271	5.708e-02	8.630e-04	7.208e-02	7.937e-01
##	2209	1.336e-01	0.1092266	4.745e-01	5.023e-01	2.258e-02	8.762e-01
##	2210	1.392e-01	0.0162475	6.187e-01	5.050e-02	9.211e-01	9.354e-02
##	2211	4.666e-01	0.2346376	6.093e-01	8.830e-01	7.444e-01	5.286e-01
##	2212	2.836e-01	0.2960749	2.696e-01	8.343e-01	8.626e-01	3.646e-01
##	2213	2.000e-01	0.4038662	8.586e-01	1.957e-01	4.887e-01	3.886e-01
##	2214	1.107e-05	0.0059194	4.996e-03	2.382e-06	7.482e-02	2.974e-03
##	2215	9.616e-01	0.8148689	9.125e-01	2.065e-01	2.601e-01	8.533e-01
##	2216	NA	NA	NA	NA	NA	NA
##	2300	8.754e-02	0.0914622	3.377e-02	3.589e-06	3.776e-06	1.196e-03
##	2301	7.824e-01	0.4706911	5.994e-02	7.877e-01	3.564e-01	6.658e-01
##	2302	9.186e-01	0.9622622	7.253e-01	4.592e-02	5.699e-03	8.712e-01
##	2303	4.356e-01	0.0158776	8.094e-05	1.798e-04	2.085e-07	8.101e-13
##	2304	9.722e-01	0.5788167	6.874e-03	1.021e-03	2.638e-01	8.714e-04
##	2305	3.865e-01	0.4570218	8.822e-01	5.863e-01	8.423e-01	7.217e-01
##	2306	8.769e-01	0.8776225	6.621e-01	4.917e-01	2.591e-02	1.254e-01
##	2307	5.442e-01	0.1882865	9.066e-02	4.054e-01	5.616e-01	1.796e-01
##	2308	2.038e-01	0.3124833	1.076e-03	1.751e-02	2.640e-02	9.101e-05
##	2309	7.839e-01	0.1244601	9.098e-01	7.595e-01	1.770e-01	3.400e-02
##	2310	2.067e-01	0.2210568	1.770e-01	3.082e-01	3.206e-01	7.593e-02
##	2311	9.501e-01	0.0210653	5.849e-01	6.183e-01	8.386e-01	3.083e-01
##	2312	5.813e-03	0.2642692	1.215e-01	1.288e-02	6.517e-02	2.792e-05
##	2400	1.185e-01	0.7168937	6.015e-01	4.727e-01	2.602e-01	5.009e-01
##	2401	NA	NA	NA	NA	NA	NA
##	2402	5.689e-01	0.4403397	1.444e-01	1.909e-02	6.933e-02	7.410e-02
##	2403	2.576e-01	0.8581028	1.454e-02	8.786e-01	6.641e-01	7.701e-03
##	2404	9.892e-01	0.2237998	3.304e-01	6.161e-01	3.819e-01	3.511e-01
##	2405	NA	NA	NA	NA	NA	NA
##	2406	7.247e-01	0.3051518	2.285e-04	1.537e-04	1.806e-04	1.485e-04
##	2500	1.291e-01	0.6134988	8.471e-01	1.753e-01	9.317e-01	9.047e-01
##	2501	NA	NA	NA	NA	NA	NA
##	2502	1.930e-02	0.0032555	7.606e-01	1.319e-01	2.799e-02	6.652e-02
##	2503	7.724e-01	0.1178199	7.816e-01	4.202e-01	7.587e-01	6.377e-01
##	2504	7.978e-01	0.7346392	1.743e-01	2.193e-01	1.461e-01	2.192e-01
##	2505	3.107e-01	0.4056434	9.559e-01	7.824e-01	8.724e-01	5.020e-01
##	2506	6.182e-01	0.8829906	4.214e-01	3.054e-01	7.285e-01	9.765e-01
##	2507	4.452e-01	0.8774375	5.595e-01	7.480e-01	6.660e-01	9.753e-01
##	2508	2.955e-01	0.5693938	7.178e-01	7.822e-01	8.379e-01	6.299e-01
##	2600	2.065e-01	0.8635877	3.501e-01	9.765e-01	4.767e-01	6.379e-04
##	2601	4.346e-02	0.9232232	3.209e-01	6.412e-01	4.981e-01	1.766e-01
##	2602	3.977e-01	0.7671064	6.711e-01	2.782e-01	1.587e-08	2.423e-01
##	2603	NA	NA	NA	NA	NA	NA
##	2604	3.007e-01	0.0624855	1.321e-01	1.401e-03	2.110e-01	1.287e-02
##	2605	NA	NA	NA	NA	NA	NA
##	2606	NA	NA	NA	NA	NA	NA
##	2607	6.344e-01	0.5820203	2.140e-01	1.573e-03	6.089e-01	2.599e-01
##	2608	NA	NA	NA	NA	NA	NA

##	2609	4.908e-01	0.0525828	9.551e-02	8.737e-04	7.218e-01	9.937e-01
##	2610	NA	NA	NA	NA	NA	NA
##	2611	2.741e-02	0.1585194	2.772e-01	3.178e-01	9.031e-01	2.094e-01
##	2612	2.810e-03	0.3302214	5.553e-01	1.897e-141	4.704e-02	8.420e-06
##	2613	1.087e-01	0.3081459	1.156e-01	7.200e-04	9.638e-02	2.413e-01
##	2614	4.356e-01	0.1367025	4.621e-01	8.612e-02	3.571e-01	4.142e-02
##	2700	7.944e-02	0.2097979	1.770e-29	3.445e-41	1.540e-45	3.466e-57
##	2701	5.742e-01	0.3153462	4.225e-03	1.073e-04	5.394e-05	3.664e-04
##	2702	5.208e-01	0.5514417	9.589e-01	3.712e-01	9.184e-01	5.720e-01
##	2703	2.576e-01	0.6486644	2.001e-01	2.533e-03	2.358e-01	2.449e-03
##	2704	5.213e-01	0.2075394	7.005e-01	5.634e-01	5.224e-01	5.997e-02
##	2705	7.917e-01	0.1013203	5.767e-01	9.097e-02	1.325e-01	3.012e-03
##	2706	3.270e-02	0.5696877	6.871e-01	1.689e-01	6.462e-01	1.702e-02
##	2707	8.414e-01	0.6514321	3.685e-01	1.197e-01	1.238e-02	3.171e-02
##	2708	3.070e-01	0.6913787	1.523e-01	4.966e-02	2.654e-01	5.374e-01
##	2709	NA	NA	NA	NA	NA	NA
##	2710	NA	NA	NA	NA	NA	NA
##	2711	4.354e-01	0.9454654	3.690e-01	2.456e-01	9.563e-02	6.054e-01
##	2712	3.555e-02	0.0739067	6.234e-01	4.000e-01	2.700e-01	1.156e-01
##	2713	3.907e-01	0.3681458	5.245e-03	3.311e-02	3.285e-04	2.086e-03
##	2714	NA	NA	NA	NA	NA	NA
##	2715	5.449e-01	0.0955916	2.870e-01	8.085e-02	9.959e-01	9.932e-01
##	2716	1.841e-01	0.1152794	1.246e-01	6.588e-01	3.000e-01	5.279e-01
##	2717	9.768e-01	0.6533929	5.952e-01	1.904e-01	2.843e-03	2.276e-04
##	2718	6.689e-01	0.1855248	4.817e-01	3.108e-01	4.377e-01	1.699e-01
##	2719	1.703e-01	0.9240962	8.590e-03	1.084e-01	2.986e-03	7.122e-04
##	2720	2.864e-01	0.7363536	9.802e-01	4.108e-01	7.667e-01	4.424e-01
##	2721	2.406e-01	0.9460884	1.675e-01	7.754e-02	8.768e-02	2.878e-01
##	2722	3.784e-01	0.5644996	5.992e-01	8.531e-01	4.578e-01	9.845e-01
##	2723	6.860e-01	0.3209558	8.844e-01	9.079e-01	2.469e-01	3.284e-02
##	2724	8.269e-01	0.2465359	6.907e-01	9.299e-01	6.787e-01	5.968e-01
##	2725	6.314e-01	0.8897737	4.556e-01	6.999e-01	2.317e-01	2.432e-01
##	2726	5.984e-01	0.3926811	1.400e-01	1.797e-01	7.649e-02	7.349e-02
##	2727	NA	NA	NA	NA	NA	NA
##	2728	1.034e-01	0.2489506	3.933e-01	4.263e-02	2.049e-01	6.174e-04
##	2729	2.595e-01	0.5065381	2.846e-02	2.812e-02	1.009e-02	1.802e-04
##	2730	7.593e-01	0.4248684	8.794e-01	4.825e-01	5.497e-01	1.717e-01
##	2731	9.721e-01	0.2049999	4.378e-02	1.851e-03	1.665e-03	7.138e-05
##	2732	4.000e-01	0.7615691	1.245e-02	1.377e-04	1.285e-03	2.701e-05
##	2733	2.544e-01	0.0451437	4.070e-02	4.557e-01	4.017e-02	3.231e-02
##	2734	2.584e-02	0.2429611	5.907e-01	8.617e-01	9.103e-01	9.882e-01
##	2735	1.221e-01	0.1950000	1.086e-02	9.834e-04	2.052e-07	5.782e-09
##	2736	5.375e-01	0.4275221	8.148e-01	9.006e-02	1.858e-03	9.843e-04
##	2737	9.788e-01	0.9606814	9.863e-01	7.944e-01	4.257e-01	2.772e-01
##	2738	2.153e-02	0.3194768	2.700e-04	4.252e-06	8.741e-08	1.504e-09
##	2739	2.123e-01	0.7237757	1.047e-12	1.155e-15	1.962e-17	3.633e-20
##	2740	1.530e-01	0.2475584	7.546e-01	9.913e-01	9.129e-01	4.759e-01
##	2741	8.725e-01	0.6940501	2.189e-01	1.739e-01	2.425e-02	1.746e-04
##	2742	2.923e-02	0.1356339	4.527e-01	8.959e-02	5.651e-01	2.353e-03
##	2743	5.546e-01	0.5530208	8.340e-02	1.286e-02	1.916e-02	1.051e-02

## 2744	NA	NA	NA	NA	NA	NA
## 2745	2.442e-03	0.0828179	3.130e-02	9.939e-01	5.933e-01	2.644e-01
## 2746	4.104e-01	0.6944815	4.845e-01	5.464e-01	8.154e-01	1.171e-01
## 2747	7.876e-01	0.3345476	6.563e-03	1.067e-03	4.944e-02	1.707e-03
## 2748	NA	NA	NA	NA	NA	NA
## 2800	6.736e-01	0.1573207	1.414e-04	5.646e-03	1.083e-02	3.239e-04
## 2801	NA	NA	NA	NA	NA	NA
## 2802	4.836e-01	0.7556317	3.679e-01	7.357e-03	8.905e-02	5.638e-03
## 2803	5.629e-01	0.9702788	7.358e-01	6.365e-01	9.322e-01	6.573e-01
## 2804	3.085e-01	0.4962781	1.996e-01	2.282e-01	1.111e-01	1.510e-01
## 2805	8.810e-01	0.0182137	9.530e-01	8.724e-03	6.959e-03	2.378e-03
## 2806	NA	NA	NA	NA	NA	NA
## 2807	NA	NA	NA	NA	NA	NA
## 2808	6.376e-02	0.3316249	5.081e-02	1.635e-03	5.825e-04	5.370e-05
## 2809	5.078e-02	0.3965838	1.546e-02	1.521e-01	1.822e-03	3.124e-03
## 2900	1.211e-02	0.2759834	9.299e-01	3.006e-02	6.702e-03	2.517e-02
## 2901	NA	NA	NA	NA	NA	NA
## 2902	7.189e-01	0.8233015	6.717e-01	4.429e-01	8.016e-02	5.014e-01
## 2903	NA	NA	NA	NA	NA	NA
## 2904	NA	NA	NA	NA	NA	NA
## 2905	NA	NA	NA	NA	NA	NA
## 2906	NA	NA	NA	NA	NA	NA
## 2907	NA	NA	NA	NA	NA	NA
## 2908	NA	NA	NA	NA	NA	NA
## 2909	NA	NA	NA	NA	NA	NA
## 2910	6.737e-02	0.1715723	8.360e-01	3.529e-05	2.197e-03	4.203e-01
## 2911	NA	NA	NA	NA	NA	NA
## 2912	NA	NA	NA	NA	NA	NA
## 2913	NA	NA	NA	NA	NA	NA
## 2914	NA	NA	NA	NA	NA	NA
## 2915	NA	NA	NA	NA	NA	NA
## 2916	5.722e-01	0.8510827	3.129e-01	2.868e-02	6.972e-03	3.469e-02
## 2917	NA	NA	NA	NA	NA	NA
## 2918	NA	NA	NA	NA	NA	NA
## 2919	NA	NA	NA	NA	NA	NA
## 2920	NA	NA	NA	NA	NA	NA
## 2921	9.228e-02	0.2408498	4.921e-01	4.015e-01	2.075e-02	2.120e-01
## 2922	NA	NA	NA	NA	NA	NA
## 2923	NA	NA	NA	NA	NA	NA
## 3000	1.474e-01	0.6635610	9.270e-01	9.394e-02	9.120e-07	3.941e-03
## 3001	NA	NA	NA	NA	NA	NA
## 3002	2.031e-01	0.2615806	5.154e-02	2.115e-01	4.371e-01	4.477e-01
## 3003	5.544e-01	0.4917048	5.064e-01	6.286e-03	1.808e-03	9.628e-04
## 3004	2.077e-01	0.4118522	2.432e-03	7.568e-04	6.686e-07	4.474e-09
## 3005	2.944e-01	0.0066293	8.322e-01	4.150e-01	3.904e-01	9.005e-01
## 3100	8.876e-01	0.8254721	9.377e-03	1.912e-01	2.322e-01	7.418e-01
## 3101	NA	NA	NA	NA	NA	NA
## 3102	NA	NA	NA	NA	NA	NA
## 3103	8.458e-01	0.8480292	1.489e-02	4.347e-01	1.124e-02	7.840e-01
## 3104	9.048e-03	0.7234097	2.659e-02	7.879e-04	4.397e-02	3.119e-01

##	3105	9.099e-02	0.0217851	5.224e-01	3.263e-01	2.272e-02	4.266e-01
##	3106	4.481e-01	0.9062182	6.245e-02	5.633e-01	9.984e-01	4.944e-01
##	3107	1.432e-01	0.6804445	4.425e-02	8.123e-02	2.867e-01	5.481e-01
##	3108	NA	NA	NA	NA	NA	NA
##	3109	3.218e-04	0.1812987	4.142e-02	8.320e-01	9.618e-01	9.296e-01
##	3110	3.292e-01	0.0984897	9.661e-01	3.612e-01	8.715e-01	8.530e-01
##	3200	5.061e-01	0.3958088	3.420e-05	1.260e-05	2.451e-04	2.571e-03
##	3201	2.577e-02	0.4425192	4.101e-01	3.823e-02	3.191e-01	6.655e-01
##	3202	7.583e-01	0.6717216	1.661e-02	2.198e-01	5.403e-05	2.021e-02
##	3203	6.712e-02	0.1080945	1.180e-03	2.047e-03	3.783e-04	3.567e-06
##	3204	3.496e-01	0.9113593	2.787e-03	1.392e-03	2.133e-04	1.014e-03
##	3205	8.730e-02	0.1687130	8.893e-01	3.156e-01	1.444e-01	3.696e-02
##	3206	5.505e-01	0.7128121	1.895e-01	2.014e-02	6.696e-02	1.406e-01
##	3207	2.934e-01	0.5600827	6.454e-02	2.119e-03	4.354e-01	2.635e-01
##	3300	9.408e-01	0.8180826	8.420e-01	1.271e-01	3.185e-01	3.786e-01
##	3301	1.094e-01	0.0075423	5.539e-02	7.505e-03	4.180e-01	1.109e-01
##	3302	5.786e-01	0.5812231	2.387e-04	2.028e-01	1.897e-01	6.799e-01
##	3303	3.220e-01	0.8787112	8.545e-02	2.298e-01	6.277e-03	2.774e-02
##	3304	9.059e-01	0.3645510	8.881e-03	3.290e-05	9.394e-03	7.595e-05
##	3305	5.945e-03	0.7684220	1.280e-03	1.407e-02	1.000e-02	1.554e-01
##	3306	1.353e-01	0.5113555	2.167e-01	7.250e-04	1.063e-02	3.471e-02
##	3307	7.541e-01	0.7074029	1.700e-01	1.872e-02	1.320e-02	1.069e-01
##	3308	3.898e-01	0.6740874	2.203e-05	2.127e-05	2.008e-03	1.629e-01
##	3309	2.196e-01	0.0405389	6.544e-02	9.026e-03	2.855e-02	7.474e-01
##	3310	9.815e-01	0.8332877	3.792e-04	7.685e-04	1.916e-01	5.535e-01
##	3311	9.143e-03	0.1909521	8.758e-01	4.810e-02	2.175e-02	1.213e-01
##	3312	3.035e-01	0.9962232	6.063e-03	1.344e-08	5.197e-03	2.367e-01
##	3313	5.614e-02	0.8166237	4.741e-01	3.199e-01	2.740e-02	2.799e-01
##	3314	7.054e-01	0.5074049	3.351e-01	3.293e-02	1.895e-01	5.515e-02
##	3315	9.104e-03	0.7552591	9.833e-02	4.098e-01	1.518e-01	5.039e-01
##	3316	5.328e-01	0.4801709	4.878e-03	4.666e-01	3.020e-01	2.219e-08
##	3317	3.276e-02	0.3912192	4.144e-02	1.566e-01	2.772e-03	1.002e-01
##	3318	8.084e-01	0.2778265	4.949e-03	8.118e-03	4.805e-02	4.829e-01
##	3319	6.251e-01	0.5000923	3.747e-02	2.679e-01	1.628e-06	4.457e-01
##	3320	3.912e-01	0.3321548	6.875e-01	6.351e-05	9.514e-01	4.016e-02
##	3321	8.980e-01	0.6685659	3.537e-01	4.128e-01	6.483e-01	6.693e-02
##	3322	9.406e-01	0.3009073	1.443e-01	1.006e-01	2.736e-03	5.356e-07
##	3400	5.863e-01	0.5301835	2.793e-01	8.316e-02	3.453e-01	1.945e-01
##	3401	NA	NA	NA	NA	NA	NA
##	3402	NA	NA	NA	NA	NA	NA
##	3403	NA	NA	NA	NA	NA	NA
##	3404	NA	NA	NA	NA	NA	NA
##	3500	9.609e-01	0.5792352	5.756e-01	6.117e-04	5.085e-02	6.465e-02
##	3501	NA	NA	NA	NA	NA	NA
##	3502	NA	NA	NA	NA	NA	NA
##	3503	NA	NA	NA	NA	NA	NA
##	3504	NA	NA	NA	NA	NA	NA
##	3505	NA	NA	NA	NA	NA	NA
##	3506	NA	NA	NA	NA	NA	NA
##	3600	NA	NA	NA	NA	NA	NA

## 3601	NA	NA	NA	NA	NA	NA
## 3602	NA	NA	NA	NA	NA	NA
## 3603	NA	NA	NA	NA	NA	NA
## 3604	NA	NA	NA	NA	NA	NA
## 3605	NA	NA	NA	NA	NA	NA
## 3606	NA	NA	NA	NA	NA	NA
## 3607	NA	NA	NA	NA	NA	NA
## 3608	NA	NA	NA	NA	NA	NA
## 3609	NA	NA	NA	NA	NA	NA
## 3610	NA	NA	NA	NA	NA	NA
## 3611	9.439e-03	0.2170453	3.795e-02	9.717e-02	1.809e-03	1.140e-02
## 3612	1.680e-02	0.3288117	2.318e-03	3.041e-04	8.103e-03	4.142e-07
## 3613	NA	NA	NA	NA	NA	NA
## 3614	NA	NA	NA	NA	NA	NA
## 3615	NA	NA	NA	NA	NA	NA
## 3616	4.969e-01	0.3636180	9.612e-01	9.121e-01	1.905e-01	8.394e-01
##	FFA2p	FLA2p	FFA3p	FLA4p		
## 1000	7.203e-01	8.567e-01	0.7545592	9.526e-01		
## 1100	1.339e-01	6.452e-01	0.1601628	9.820e-01		
## 1101	4.304e-02	7.678e-01	0.0427941	5.350e-01		
## 1102	8.067e-01	6.964e-01	0.8384494	7.166e-01		
## 1103	8.753e-01	6.571e-01	0.7901681	6.331e-01		
## 1104	6.269e-01	8.453e-01	0.6479708	9.222e-01		
## 1105	9.277e-01	2.056e-01	0.8914660	2.075e-01		
## 1106	8.265e-04	4.003e-01	0.0010033	5.740e-01		
## 1107	3.223e-01	3.041e-01	0.2115410	2.173e-01		
## 1108	7.839e-02	8.332e-01	0.0688806	6.981e-01		
## 1109	3.884e-01	4.299e-02	0.8426587	5.888e-02		
## 1110	7.312e-01	5.593e-01	0.6731997	5.282e-01		
## 1111	3.213e-01	9.322e-01	0.3291653	9.904e-01		
## 1200	NA	NA	NA	NA		
## 1201	8.204e-01	5.193e-01	0.3721549	2.737e-01		
## 1202	4.323e-01	2.867e-01	0.6461531	3.779e-01		
## 1203	5.379e-01	8.121e-01	0.1755215	2.045e-01		
## 1204	2.825e-01	6.902e-01	0.3029414	7.685e-01		
## 1205	NA	NA	NA	NA		
## 1206	NA	NA	NA	NA		
## 1207	1.698e-02	8.205e-01	0.0003163	1.099e-02		
## 1208	1.724e-01	1.508e-01	0.9001383	4.793e-01		
## 1209	NA	NA	NA	NA		
## 1210	2.938e-01	6.491e-01	0.0428712	7.795e-02		
## 1211	5.401e-01	4.448e-01	0.5185838	3.673e-01		
## 1212	2.789e-01	3.364e-01	0.8261028	9.848e-01		
## 1213	9.913e-01	8.704e-01	0.7389809	7.060e-01		
## 1300	2.935e-02	3.702e-02	0.1067254	1.348e-01		
## 1301	NA	NA	NA	NA		
## 1302	NA	NA	NA	NA		
## 1303	9.110e-02	6.347e-01	0.0976575	7.252e-01		
## 1304	1.187e-01	7.477e-02	0.1907937	1.053e-01		
## 1305	9.980e-01	2.435e-01	0.9009675	2.402e-01		

```
## 1306 7.601e-01 3.067e-01 0.6628905 2.692e-01
## 1307 1.883e-01 7.706e-01 0.1878912 7.827e-01
## 1308 2.058e-01 4.558e-01 0.1577032 3.115e-01
## 1309 4.746e-01 2.623e-01 0.5669819 3.016e-01
## 1310 5.799e-01 5.809e-01 0.5938890 5.947e-01
## 1311 4.731e-01 6.934e-01 0.4506597 6.441e-01
## 1312 4.479e-01 2.310e-01 0.4935884 2.494e-01
## 1313 8.493e-01 7.807e-02 0.6988575 7.803e-02
## 1314 2.040e-01 3.001e-01 0.1714894 2.358e-01
## 1315 2.042e-01 6.372e-02 0.3583668 1.007e-01
## 1400 8.009e-01 4.096e-03 0.0224685 1.356e-04
## 1401 2.150e-01 3.993e-01 0.3500229 7.810e-01
## 1402 7.164e-01 4.225e-01 0.9687956 4.905e-01
## 1403 8.155e-01 6.277e-01 0.9199641 6.396e-01
## 1404 1.493e-01 2.806e-01 0.2330000 6.593e-01
## 1405 7.267e-01 3.370e-01 0.6368549 2.975e-01
## 1406 5.543e-02 1.143e-02 0.3420082 1.084e-01
## 1407 9.197e-01 5.647e-02 0.0754018 1.151e-02
## 1408 7.305e-01 6.590e-01 0.9191116 7.653e-01
## 1409 5.904e-02 6.410e-01 0.0158685 9.798e-02
## 1410 9.612e-01 7.432e-01 0.7383167 7.123e-01
## 1500 8.904e-01 6.903e-03 0.7086081 7.078e-03
## 1501 8.328e-03 8.688e-01 0.0076446 9.344e-01
## 1502 4.493e-01 8.048e-04 0.6042604 1.213e-03
## 1503 1.269e-01 1.521e-01 0.0920167 1.187e-01
## 1504      NA      NA      NA      NA
## 1505 3.994e-02 9.076e-01 0.0905037 9.374e-01
## 1506      NA      NA      NA      NA
## 1507      NA      NA      NA      NA
## 1508      NA      NA      NA      NA
## 1600 1.474e-01 6.115e-01 0.1319945 5.110e-01
## 1601      NA      NA      NA      NA
## 1602 4.964e-01 7.776e-01 0.5007694 8.044e-01
## 1603 2.221e-01 4.962e-01 0.2798131 7.491e-01
## 1604 9.376e-01 6.798e-01 0.9103112 6.759e-01
## 1605 2.048e-01 5.533e-01 0.2080523 5.625e-01
## 1606 3.780e-02 3.983e-01 0.0370325 3.877e-01
## 1607 1.725e-01 3.459e-03 0.0485817 8.054e-04
## 1700 2.607e-02 3.200e-01 0.0549718 9.447e-01
## 1701      NA      NA      NA      NA
## 1702 2.740e-01 6.071e-01 0.0532513 5.674e-02
## 1703 9.961e-01 2.516e-01 0.6519312 1.830e-01
## 1704      NA      NA      NA      NA
## 1705 8.891e-01 6.914e-01 0.9381553 6.618e-01
## 1706 1.430e-01 1.286e-01 0.3475250 3.002e-01
## 1707 1.978e-01 7.593e-07 0.8199626 0.000e+00
## 1708 5.028e-01 6.233e-01 0.6169174 7.180e-01
## 1709 3.019e-01 3.039e-01 0.9697288 3.934e-01
## 1710 4.697e-01 8.643e-01 0.3693475 5.630e-01
## 1711 1.221e-02 1.454e-01      NA 1.722e-01
```

##	1712	4.990e-02	2.358e-01	0.0826093	3.770e-01
##	1800	NA	NA	NA	NA
##	1801	NA	NA	NA	NA
##	1802	8.948e-01	6.547e-01	0.8739099	6.677e-01
##	1803	6.863e-01	3.601e-01	0.3100184	2.089e-01
##	1804	5.976e-01	9.811e-01	0.5762371	8.843e-01
##	1900	8.890e-02	2.165e-01	0.2908705	8.363e-01
##	1901	4.775e-01	8.041e-01	0.4533407	6.892e-01
##	1902	8.863e-01	3.202e-01	0.7932128	3.843e-01
##	1903	NA	NA	NA	NA
##	1904	9.433e-03	1.348e-02	0.1208197	2.737e-01
##	1905	NA	NA	NA	NA
##	1906	8.902e-02	5.196e-01	0.0996796	7.194e-01
##	1907	3.938e-01	9.689e-01	0.3816744	7.784e-01
##	1908	7.555e-01	7.743e-01	0.7919796	8.157e-01
##	1909	4.448e-01	3.249e-01	0.2381772	1.742e-01
##	1910	3.325e-01	5.704e-01	NA	3.342e-01
##	1911	4.425e-01	5.963e-01	0.5388459	8.261e-01
##	1912	5.048e-02	4.835e-01	0.0469994	5.835e-01
##	1913	3.822e-01	3.639e-01	0.4224557	3.574e-01
##	2000	2.466e-01	2.949e-02	0.0104880	2.093e-03
##	2001	4.191e-04	2.005e-02	0.0732560	7.907e-01
##	2002	6.887e-01	5.234e-01	0.3107672	2.539e-01
##	2003	6.189e-01	7.469e-01	0.7074716	8.967e-01
##	2100	7.446e-01	9.802e-01	0.7607504	9.895e-01
##	2101	NA	NA	NA	NA
##	2102	7.585e-02	3.561e-01	0.0554986	6.323e-01
##	2103	9.452e-01	7.739e-03	0.6001975	6.678e-05
##	2104	NA	NA	NA	NA
##	2105	3.821e-01	1.850e-02	0.6957791	2.767e-02
##	2200	2.042e-01	4.265e-01	0.4298832	8.110e-01
##	2201	7.945e-01	7.994e-01	0.5701260	6.035e-01
##	2202	5.845e-01	5.001e-01	0.2649916	3.058e-01
##	2203	NA	NA	NA	NA
##	2204	5.510e-01	1.058e-01	0.8447194	1.276e-01
##	2205	9.924e-01	9.920e-01	0.9862575	9.873e-01
##	2206	NA	NA	NA	NA
##	2207	2.781e-03	3.773e-02	0.0127094	3.813e-01
##	2208	1.674e-01	1.415e-04	0.8505230	8.758e-04
##	2209	1.831e-01	1.828e-01	0.3022088	3.866e-01
##	2210	1.350e-01	3.595e-02	0.5205932	4.816e-02
##	2211	3.476e-01	1.749e-01	0.6938191	2.368e-01
##	2212	1.237e-01	2.575e-01	0.1519446	9.161e-01
##	2213	3.158e-01	3.025e-01	0.1838036	1.629e-01
##	2214	3.321e-01	5.631e-01	0.2576583	3.664e-01
##	2215	5.607e-01	2.904e-01	0.8958894	5.461e-01
##	2216	NA	NA	NA	NA
##	2300	6.236e-02	1.388e-01	0.1516743	4.989e-01
##	2301	8.712e-01	3.480e-01	0.6869238	2.987e-01
##	2302	6.174e-01	1.600e-01	0.9512587	1.590e-01

```
## 2303 3.535e-01 4.566e-02 0.5451449 6.208e-02
## 2304 8.167e-01 6.691e-01 0.8009919 6.646e-01
## 2305 4.364e-01 4.676e-01 0.5682845 6.496e-01
## 2306 8.113e-01 9.730e-01 0.8052012 9.867e-01
## 2307 4.091e-01 9.257e-02 0.6699552 1.557e-01
## 2308 9.643e-02 2.348e-01 0.2381249 6.364e-01
## 2309 7.179e-01 1.436e-01 0.9798157 1.622e-01
## 2310 2.561e-01 1.989e-01 0.2520127 1.973e-01
## 2311 9.453e-01 9.802e-03 0.5816505 1.147e-02
## 2312 9.228e-03 1.678e-01 0.0222915 4.678e-01
## 2400 1.072e-01 6.457e-01 0.0928564 5.516e-01
## 2401      NA      NA      NA      NA
## 2402 9.243e-01 6.498e-01 0.9629063 6.495e-01
## 2403 1.908e-01 6.820e-01 0.1712632 5.706e-01
## 2404 8.872e-01 2.180e-01 0.8160305 2.127e-01
## 2405      NA      NA 0.7653062      NA
## 2406 2.710e-01 4.599e-01 0.2297737 4.024e-01
## 2500 1.343e-01 6.186e-01 0.1120586 4.736e-01
## 2501      NA      NA      NA      NA
## 2502 1.923e-02 2.162e-02 0.0327092 5.704e-02
## 2503 8.363e-01 1.910e-01 0.9317390 1.985e-01
## 2504 8.998e-01 7.211e-01 0.9634599 7.364e-01
## 2505 3.399e-01 4.159e-01 0.3274904 4.009e-01
## 2506 2.863e-01 7.003e-01 0.3138336 8.330e-01
## 2507 3.873e-01 8.414e-01 0.3884508 8.988e-01
## 2508 3.238e-01 5.514e-01 0.2901355 4.896e-01
## 2600 4.044e-01 3.939e-01 0.0629301 5.654e-02
## 2601 1.018e-01 9.349e-01 0.1171209 5.356e-01
## 2602 6.967e-01 3.087e-01 0.2860563 5.715e-04
## 2603      NA      NA      NA      NA
## 2604 6.553e-02 1.608e-01 0.1606176 3.669e-01
## 2605 2.360e-01 7.811e-01 0.1989899 5.723e-01
## 2606      NA      NA      NA      NA
## 2607 7.820e-01 1.582e-01 0.9612716 1.276e-01
## 2608      NA      NA 0.3239950 3.180e-01
## 2609 2.476e-01 1.571e-03 0.3608330 5.650e-02
## 2610 4.404e-06 1.576e-05 0.0004869 6.422e-01
## 2611 6.245e-03 3.756e-02 0.0160651 2.640e-01
## 2612 4.015e-02 2.488e-01 0.0391112 3.738e-01
## 2613 5.640e-02 1.645e-01 0.0314770 6.914e-02
## 2614 3.965e-01 7.725e-02 0.8372891 1.403e-01
## 2700 1.282e-03 3.952e-03 0.0167316 4.239e-02
## 2701 2.222e-01 3.413e-01 0.1465241 2.006e-01
## 2702 6.006e-01 3.617e-01 0.4356074 2.845e-01
## 2703 3.795e-01 3.688e-01 0.1743667 1.533e-01
## 2704 9.400e-01 5.115e-01 0.9421066 5.052e-01
## 2705 5.026e-01 1.488e-01 0.4181607 1.307e-01
## 2706 2.771e-01 8.420e-01 0.2543248 7.870e-01
## 2707 8.966e-01 9.214e-01 0.8878405 9.193e-01
## 2708 1.757e-01 4.966e-01 0.1377873 3.229e-01
```

## 2709	NA	NA	NA	NA
## 2710	7.589e-01	2.963e-01	0.6301669	2.332e-01
## 2711	5.220e-01	8.440e-01	0.5235213	7.890e-01
## 2712	2.221e-02	7.087e-02	0.0460819	1.664e-01
## 2713	5.964e-01	9.968e-01	0.5910957	9.963e-01
## 2714	1.558e-02	9.516e-01	0.0249342	3.185e-01
## 2715	3.259e-01	1.479e-01	0.4775386	1.773e-01
## 2716	2.082e-01	5.513e-02	0.0771357	1.796e-02
## 2717	7.995e-01	8.423e-01	0.8275571	8.767e-01
## 2718	7.067e-01	1.374e-01	0.9076197	1.582e-01
## 2719	1.311e-01	7.472e-01	0.1537963	7.821e-01
## 2720	3.150e-01	6.084e-01	0.3372748	6.602e-01
## 2721	2.589e-01	9.857e-01	0.2447004	9.974e-01
## 2722	3.539e-01	7.078e-01	0.4093702	7.676e-01
## 2723	5.224e-01	3.723e-01	0.5756603	4.225e-01
## 2724	7.484e-01	2.415e-01	0.8361099	2.381e-01
## 2725	8.469e-01	5.978e-01	0.8553958	6.058e-01
## 2726	8.995e-01	2.583e-01	0.9855349	2.614e-01
## 2727	2.049e-01	9.938e-01	0.2093799	8.819e-01
## 2728	1.085e-01	3.535e-01	0.0831551	2.509e-01
## 2729	3.896e-02	5.841e-01	0.0463652	8.763e-01
## 2730	4.655e-01	4.752e-01	0.4112424	4.088e-01
## 2731	1.700e-01	4.452e-01	0.1574761	3.959e-01
## 2732	3.880e-01	4.576e-01	0.2486620	2.850e-01
## 2733	4.620e-01	2.255e-01	0.6312608	3.076e-01
## 2734	2.141e-02	2.162e-01	0.0407929	8.274e-01
## 2735	4.082e-01	3.077e-01	0.4791729	3.613e-01
## 2736	4.753e-01	8.295e-02	0.9969047	1.111e-01
## 2737	6.837e-01	8.796e-01	0.7014671	9.543e-01
## 2738	6.769e-03	4.980e-01	0.0047843	2.443e-01
## 2739	6.289e-01	3.268e-01	0.4433708	2.611e-01
## 2740	1.188e-01	3.145e-01	0.1334060	3.874e-01
## 2741	8.549e-01	7.886e-01	0.8010974	7.334e-01
## 2742	2.525e-02	1.195e-01	0.0027915	1.237e-02
## 2743	2.620e-01	8.931e-01	0.2572242	9.265e-01
## 2744	NA	NA	NA	NA
## 2745	6.299e-06	9.546e-04	0.0005171	1.244e-01
## 2746	2.134e-01	8.601e-01	0.2127970	8.402e-01
## 2747	7.740e-01	1.526e-01	0.7350552	1.521e-01
## 2748	6.489e-01	8.286e-01	0.6134537	7.471e-01
## 2800	9.526e-02	2.002e-01	0.1199430	2.807e-01
## 2801	NA	NA	NA	NA
## 2802	8.356e-01	9.671e-01	0.8296431	9.541e-01
## 2803	6.660e-01	6.034e-01	0.5945291	5.645e-01
## 2804	1.711e-01	4.372e-01	0.1828205	5.296e-01
## 2805	7.264e-01	3.330e-02	0.6870270	3.734e-02
## 2806	9.294e-01	9.995e-01	0.9295148	9.970e-01
## 2807	NA	NA	NA	NA
## 2808	1.158e-02	4.823e-01	0.0091326	4.831e-01
## 2809	2.233e-01	5.439e-01	0.1976317	4.566e-01

##	2900	5.365e-02	5.624e-02	0.0012030	1.060e-03
##	2901	NA	NA	NA	NA
##	2902	9.996e-01	6.835e-01	0.8110838	6.089e-01
##	2903	NA	NA	NA	NA
##	2904	NA	NA	NA	NA
##	2905	8.657e-01	1.897e-01	0.7757714	1.769e-01
##	2906	NA	NA	NA	NA
##	2907	NA	NA	NA	NA
##	2908	NA	NA	NA	NA
##	2909	NA	NA	NA	NA
##	2910	2.191e-01	4.887e-01	0.3556920	5.995e-01
##	2911	NA	NA	NA	NA
##	2912	NA	NA	NA	NA
##	2913	NA	NA	NA	NA
##	2914	NA	NA	NA	NA
##	2915	NA	NA	NA	NA
##	2916	7.436e-01	7.898e-01	0.7919649	8.382e-01
##	2917	NA	NA	NA	NA
##	2918	NA	NA	NA	NA
##	2919	NA	NA	NA	NA
##	2920	NA	NA	NA	NA
##	2921	3.107e-02	3.973e-01	0.0316239	2.301e-01
##	2922	NA	NA	NA	NA
##	2923	NA	NA	NA	NA
##	3000	8.329e-02	4.059e-02	0.0104775	4.774e-03
##	3001	NA	NA	NA	NA
##	3002	1.093e-01	8.170e-01	0.0987328	6.936e-01
##	3003	5.816e-01	7.012e-01	0.5016509	5.959e-01
##	3004	5.099e-02	3.546e-02	0.2354645	9.953e-02
##	3005	2.141e-01	1.022e-02	0.0673307	3.855e-03
##	3100	4.910e-01	7.091e-01	0.4809370	6.717e-01
##	3101	5.137e-02	2.719e-01	0.1528391	5.950e-01
##	3102	NA	NA	0.8528091	8.528e-01
##	3103	5.303e-01	8.487e-01	0.4983614	7.229e-01
##	3104	1.343e-02	7.300e-01	0.0126222	8.338e-01
##	3105	2.813e-01	4.286e-02	0.8705942	2.321e-02
##	3106	6.723e-01	8.239e-01	0.6709894	8.347e-01
##	3107	1.489e-01	7.992e-01	0.1315732	9.380e-01
##	3108	NA	NA	NA	NA
##	3109	2.317e-03	9.296e-02	NA	9.645e-01
##	3110	3.052e-01	4.105e-02	0.3521949	4.290e-02
##	3200	5.815e-01	3.993e-01	NA	5.006e-01
##	3201	9.140e-03	2.344e-01	0.0343444	8.465e-01
##	3202	9.454e-01	4.411e-01	0.8254874	4.049e-01
##	3203	3.401e-01	2.578e-02	0.2283119	2.201e-02
##	3204	4.906e-01	7.181e-01	0.3726482	4.780e-01
##	3205	1.329e-01	2.164e-01	0.2337115	4.416e-01
##	3206	6.500e-02	7.972e-01	0.0623552	6.986e-01
##	3207	3.756e-01	4.909e-01	0.1312247	1.797e-01
##	3300	9.112e-01	8.576e-01	0.9789734	8.836e-01

##	3301	2.214e-01	2.181e-02	0.7029552	5.868e-02
##	3302	1.527e-01	6.335e-01	0.3045676	7.563e-01
##	3303	4.074e-01	8.981e-01	0.2781309	4.262e-01
##	3304	9.400e-01	4.735e-01	0.4745019	3.181e-01
##	3305	3.246e-03	6.134e-01	0.0013492	6.979e-02
##	3306	9.629e-02	1.163e-01	0.2813297	3.368e-01
##	3307	7.216e-01	7.480e-01	0.7865932	8.160e-01
##	3308	1.537e-01	7.328e-01	0.1350082	4.442e-01
##	3309	3.919e-01	1.034e-01	0.8825456	2.180e-01
##	3310	7.588e-01	7.320e-01	0.3670133	3.597e-01
##	3311	1.017e-01	4.699e-01	0.0132002	5.631e-02
##	3312	4.251e-02	5.210e-01	0.0264414	2.382e-01
##	3313	5.268e-02	6.622e-01	0.0304184	3.040e-01
##	3314	5.949e-01	6.418e-01	0.1534428	1.601e-01
##	3315	2.192e-02	7.938e-01	0.0003421	3.484e-03
##	3316	6.617e-01	6.891e-01	0.8536462	9.733e-01
##	3317	3.832e-03	1.636e-01	0.0029346	2.545e-02
##	3318	7.346e-01	6.719e-01	0.9568280	8.133e-01
##	3319	5.708e-01	4.622e-01	0.3051826	2.720e-01
##	3320	4.695e-01	3.223e-01	0.8715410	4.949e-01
##	3321	7.537e-01	3.847e-01	0.7503710	3.958e-01
##	3322	8.293e-01	3.405e-01	0.4940277	2.597e-01
##	3400	4.820e-01	5.763e-01	0.4464283	5.444e-01
##	3401	NA	NA	NA	NA
##	3402	NA	NA	NA	NA
##	3403	NA	NA	NA	NA
##	3404	NA	NA	NA	NA
##	3500	6.883e-01	8.379e-01	0.5539364	6.227e-01
##	3501	NA	NA	NA	NA
##	3502	NA	NA	NA	NA
##	3503	NA	NA	NA	NA
##	3504	1.774e-01	3.120e-01	0.1000774	5.479e-01
##	3505	NA	NA	NA	NA
##	3506	NA	NA	NA	NA
##	3600	2.061e-01	2.398e-01	0.4880175	5.314e-01
##	3601	NA	NA	NA	NA
##	3602	NA	NA	NA	NA
##	3603	NA	NA	NA	NA
##	3604	NA	NA	NA	NA
##	3605	NA	NA	NA	NA
##	3606	NA	NA	NA	NA
##	3607	NA	NA	NA	NA
##	3608	NA	NA	NA	NA
##	3609	NA	NA	NA	NA
##	3610	NA	NA	NA	NA
##	3611	9.205e-02	4.445e-01	0.1003324	5.576e-01
##	3612	4.056e-02	1.055e-01	0.0048420	1.319e-02
##	3613	NA	NA	NA	NA
##	3614	NA	NA	NA	NA

```
## 3615          NA          NA          NA          NA
## 3616 3.001e-01 4.160e-01 0.4668213 6.457e-01
```

```
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	"**"	"*"	" "	" "	" "	" "	"**"	"*"	" "

[illegible]

##	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 "" "" "" "" "" "" "" "" "" ""

```

```
print("Gender-based average team sizes in 2018")
```

```
## [1] "Gender-based average team sizes in 2018"
```

```
print(GenderAnalysed)
```

```

##      CitationSet Female1st FemaleLast
## 1000          43         15         12
## 1100          61         35         17
## 1101           8          3          3

```


## 1102	77	28	31
## 1103	87	38	32
## 1104	69	29	15
## 1105	138	56	42
## 1106	62	30	21
## 1107	15	2	4
## 1108	31	15	14
## 1109	44	22	18
## 1110	64	27	21
## 1111	29	10	9
## 1200	4	1	3
## 1201	43	25	25
## 1202	60	27	23
## 1203	37	20	23
## 1204	9	2	1
## 1205	1	1	0
## 1206	5	5	0
## 1207	24	11	12
## 1208	18	10	8
## 1209	3	3	0
## 1210	5	3	2
## 1211	39	13	10
## 1212	17	4	3
## 1213	31	20	17
## 1300	56	32	17
## 1301	1	0	0
## 1302	6	5	2
## 1303	49	23	21
## 1304	18	10	1
## 1305	38	17	10
## 1306	8	3	4
## 1307	23	13	10
## 1308	18	12	5
## 1309	7	5	3
## 1310	20	15	5
## 1311	40	16	11
## 1312	44	20	17
## 1313	15	7	5
## 1314	34	17	11
## 1315	5	2	2
## 1400	11	3	2
## 1401	10	4	4
## 1402	27	9	10
## 1403	31	12	11
## 1404	10	3	6
## 1405	18	4	5
## 1406	28	15	14
## 1407	18	6	6
## 1408	44	19	22
## 1409	46	22	26

## 1410	4	1	0
## 1500	19	4	2
## 1501	1	0	0
## 1502	9	3	1
## 1503	12	2	2
## 1504	0	0	0
## 1505	2	0	0
## 1506	1	0	0
## 1507	5	1	0
## 1508	5	1	0
## 1600	36	10	6
## 1601	5	4	2
## 1602	16	4	3
## 1603	2	2	0
## 1604	9	3	4
## 1605	20	8	7
## 1606	24	7	7
## 1607	4	1	0
## 1700	9	1	1
## 1701	1	0	0
## 1702	4	1	2
## 1703	2	0	0
## 1704	3	0	0
## 1705	5	1	0
## 1706	15	2	2
## 1707	4	0	0
## 1708	2	2	0
## 1709	5	2	1
## 1710	13	4	2
## 1711	5	1	0
## 1712	11	1	0
## 1800	3	1	1
## 1801	0	0	0
## 1802	7	2	3
## 1803	9	1	2
## 1804	2	0	0
## 1900	20	7	8
## 1901	27	12	4
## 1902	12	2	1
## 1903	4	0	0
## 1904	30	7	7
## 1905	2	0	0
## 1906	20	9	6
## 1907	42	21	4
## 1908	40	16	8
## 1909	25	8	1
## 1910	16	6	6
## 1911	9	2	3
## 1912	5	4	1
## 1913	3	2	0

## 2000	20	1	6
## 2001	10	1	4
## 2002	58	13	14
## 2003	38	10	11
## 2100	6	3	1
## 2101	1	0	0
## 2102	8	2	1
## 2103	5	2	1
## 2104	4	2	1
## 2105	28	9	8
## 2200	9	1	1
## 2201	3	1	0
## 2202	3	0	0
## 2203	2	1	1
## 2204	18	5	4
## 2205	28	8	2
## 2206	0	0	0
## 2207	2	0	0
## 2208	8	1	2
## 2209	16	3	3
## 2210	23	4	2
## 2211	22	5	2
## 2212	5	1	0
## 2213	18	5	4
## 2214	3	0	0
## 2215	21	6	1
## 2216	5	3	1
## 2300	33	13	12
## 2301	10	3	1
## 2302	10	3	2
## 2303	97	36	22
## 2304	26	13	5
## 2305	19	8	3
## 2306	11	5	2
## 2307	18	14	8
## 2308	51	21	23
## 2309	33	12	9
## 2310	30	16	5
## 2311	21	9	3
## 2312	59	19	14
## 2400	11	4	4
## 2401	1	0	0
## 2402	7	5	2
## 2403	13	7	2
## 2404	19	9	5
## 2405	11	6	2
## 2406	6	5	4
## 2500	30	7	0
## 2501	0	0	0
## 2502	10	5	1

## 2503	4	1	2
## 2504	17	2	1
## 2505	10	1	1
## 2506	9	2	0
## 2507	7	3	3
## 2508	15	5	3
## 2600	20	4	4
## 2601	1	0	0
## 2602	8	0	0
## 2603	3	0	0
## 2604	12	2	3
## 2605	2	1	0
## 2606	4	1	1
## 2607	5	0	0
## 2608	6	0	0
## 2609	6	0	0
## 2610	4	1	0
## 2611	7	2	3
## 2612	4	0	0
## 2613	8	0	0
## 2614	9	0	0
## 2700	100	48	40
## 2701	34	22	10
## 2702	5	2	3
## 2703	11	5	2
## 2704	5	2	0
## 2705	22	8	3
## 2706	6	4	2
## 2707	8	3	2
## 2708	9	4	4
## 2709	0	0	0
## 2710	2	1	2
## 2711	8	5	4
## 2712	18	9	8
## 2713	9	6	5
## 2714	7	1	2
## 2715	8	5	3
## 2716	5	2	1
## 2717	21	15	8
## 2718	6	3	1
## 2719	36	15	15
## 2720	7	4	2
## 2721	2	1	0
## 2722	5	2	3
## 2723	15	5	3
## 2724	15	5	5
## 2725	23	13	7
## 2726	17	6	5
## 2727	2	1	1
## 2728	29	13	8

## 2729	21	19	15
## 2730	18	10	8
## 2731	24	8	9
## 2732	48	18	7
## 2733	17	9	9
## 2734	16	8	7
## 2735	43	24	18
## 2736	110	74	68
## 2737	22	9	4
## 2738	64	41	28
## 2739	116	76	60
## 2740	13	4	2
## 2741	19	7	4
## 2742	13	9	4
## 2743	11	10	5
## 2744	0	0	0
## 2745	5	2	1
## 2746	55	13	11
## 2747	4	0	0
## 2748	4	1	1
## 2800	31	14	6
## 2801	4	3	0
## 2802	25	14	7
## 2803	6	2	1
## 2804	12	7	4
## 2805	10	4	4
## 2806	1	1	0
## 2807	2	2	0
## 2808	18	6	4
## 2809	9	5	1
## 2900	31	26	28
## 2901	1	1	0
## 2902	4	2	2
## 2903	0	0	0
## 2904	0	0	0
## 2905	3	2	1
## 2906	2	1	1
## 2907	1	1	0
## 2908	0	0	0
## 2909	8	7	4
## 2910	9	3	5
## 2911	1	1	0
## 2912	0	0	0
## 2913	2	2	0
## 2914	1	0	0
## 2915	0	0	0
## 2916	44	30	21
## 2917	1	1	0
## 2918	0	0	0
## 2919	1	1	0

## 2920	1	1	0
## 2921	11	8	4
## 2922	0	0	0
## 2923	0	0	0
## 3000	3	1	0
## 3001	1	0	0
## 3002	14	7	5
## 3003	26	11	11
## 3004	38	23	13
## 3005	13	11	5
## 3100	10	0	0
## 3101	6	2	0
## 3102	4	1	1
## 3103	2	2	0
## 3104	25	4	2
## 3105	13	4	3
## 3106	1	0	0
## 3107	22	3	3
## 3108	1	1	0
## 3109	5	1	0
## 3110	6	2	1
## 3200	27	12	8
## 3201	4	2	1
## 3202	24	16	13
## 3203	22	15	13
## 3204	48	33	31
## 3205	33	17	12
## 3206	4	2	1
## 3207	36	22	21
## 3300	30	13	14
## 3301	37	23	23
## 3302	8	2	1
## 3303	31	17	14
## 3304	211	142	131
## 3305	100	52	52
## 3306	42	24	19
## 3307	4	1	1
## 3308	47	21	19
## 3309	11	4	3
## 3310	47	29	31
## 3311	16	7	2
## 3312	113	61	58
## 3313	16	6	5
## 3314	35	22	22
## 3315	35	13	13
## 3316	68	42	42
## 3317	9	4	5
## 3318	19	17	18
## 3319	12	10	9
## 3320	37	10	12

## 3321	12	5	5
## 3322	11	5	2
## 3400	26	9	6
## 3401	0	0	0
## 3402	1	1	0
## 3403	5	3	0
## 3404	2	2	0
## 3500	7	2	1
## 3501	0	0	0
## 3502	0	0	0
## 3503	0	0	0
## 3504	2	0	0
## 3505	2	1	0
## 3506	0	0	0
## 3600	0	0	0
## 3601	2	2	0
## 3602	3	2	2
## 3603	2	1	0
## 3604	1	1	0
## 3605	1	0	0
## 3606	0	0	0
## 3607	1	0	0
## 3608	0	0	0
## 3609	1	1	0
## 3610	7	3	4
## 3611	9	6	6
## 3612	42	22	9
## 3613	0	0	0
## 3614	4	1	1
## 3615	0	0	0
## 3616	13	8	7

```

write.table(BasicStats, file=paste("D:\\Downloads\\basicStats", Country,
MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3b.txt", sep=""))
write.table(BasicStats2, file=paste("D:\\Downloads\\basicStats2", Country,
MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3b.txt", sep=""))
write.table(GenderTeamSize, file=paste("D:\\Downloads\\GenderTeamSize",
Country, MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3b.txt", sep=""))
write.table(RegCoef, file=paste("D:\\Downloads\\RegCoef", Country,
MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3b.txt", sep=""))
write.table(RegP, file=paste("D:\\Downloads\\RegP", Country,
MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3b.txt", sep=""))
write.table(RegStar, file=paste("D:\\Downloads\\RegStar", Country,
MaxCountriesAllowed, " ", FirstYearForCitationAnalysis, "-",
LastYearForCitationAnalysis, "var3b.txt", sep=""))

```